Online moving horizon estimation of fluxes in metabolic reaction networks

D. Vercammen, F. Logist, J. Van Impe*

KU Leuven, BioTeC - Chemical and Biochemical Process Technology and Control & OPTEC - Center of Excellence: Optimization in Engineering, Department of Chemical Engineering, Gebroeders De Smetstraat 1, 9000 Gent, Belgium

Abstract

Using online state and parameter estimation, concentrations and fluxes in bioprocesses can be estimated for use in monitoring, optimization and control applications. Existing methodologies, however, either ignore the dynamic nature of the problem, or focus on the extracellular concentration states and pay less attention to accurate flux estimates. These estimates are useful for online monitoring of the flux state of an organism, or for developing novel flux-based strategies for online control of bioreactors.

In this contribution, the dynamic metabolic flux analysis model structure is combined with two kinetic flux models: a linear flux model and a nonlinear, more mechanistic flux model. The parameters of these models are estimated online through a moving horizon estimation strategy. The resulting algorithm is illustrated on two simulated case studies: a small-scale network, to assess the influence of important algorithm parameters on the final estimates, and a medium-scale network for *Escherichia coli*, to empiri-

*Corresponding author. E-mail: jan.vanimpe@cit.kuleuven.be, Tel.: +32 16 32 14 66

Postprint of manuscript accepted for Journal of Process Control

cally test the performance of the methodology in a more realistic situation.

An important parameter in this estimation strategy is the chosen noise level on the estimated parameters. This choice is not trivial, but is observed to have a significant influence on the resulting estimates. Furthermore, also the effect of the choice of the null space basis for the stoichiometric matrix of the metabolic reaction network was assessed. In the small-scale case study, it was found that a linear flux model with a specific parameter noise level was performing well for both state and flux estimation. The influence of the choice of the null space basis matrix on the estimation performance was much lower. The resulting scenario was evaluated in the medium-scale case study and found to be performing very well also in that case.

Key words:
moving horizon flux estimation, dynamic metabolic flux analysis, online state and parameter estimation, (non)linear kinetic flux models

1. Introduction

State and parameter estimation in bioprocesses is used to estimate unmeasured states, and/or unknown parameters from possibly noisy data from the output of the biosystem. To this end, a bioprocess model is used to relate inputs, outputs and states to each other. The estimates can be used for monitoring of the bioprocess, to get a better understanding of what is going on at that point in time in the bioprocess, as a soft-sensor for variables that cannot be measured, and as a basis for model-based predictive control of the bioprocess [1, 2].

In most of these monitoring and control applications, simple, unstructured models are used as the basis for state and parameter estimation (e.g., [3, 4, 5]), or focus is put on the estimation of the specific growth rate [6, 7, 8]. Methods for estimation based on structured models, and more specifically metabolic reaction network-based models like, e.g.,(dynamic) metabolic flux analysis (MFA/dMFA) [9, 10], are less widespread. These methodologies use a pseudo steady-state assumption to provide an experimentally and computationally tractable (dynamic) model structure. Applications in literature can generally be divided into two approaches, which both use online sampling: methods that combine an online flux calculation scheme with a classical MFA approach in which a static MFA problem is solved at every sampling instant, and methods that combine a dMFA type model with classic state estimation techniques like Kalman filtering, e.g., the extended Kalman filter (EKF), or moving horizon estimation (MHE).

Examples of the first class can be found in, e.g., [11] and [12]. Dias et al.[11] use an online adaptive MFA framework to estimate flux distributions during polyhydroxybutyrate production in a mixed microbial culture. In [12], real-time metabolic flux analysis is performed during experiments with *Pichia pastoris*, using near infrared spectroscopy for measuring the exchange fluxes that are used as input to the MFA problems. These methods are focused on getting accurate estimates for the fluxes, which is important, as knowledge of the intracellular fluxes over time can significantly enhance the knowledge gained during the monitoring of bioprocesses, and help in choos-

Journal homepage: http://www.journals.elsevier.com/journal-of-process-control
Original file available at: doi:10.1016/j.jprocont.2015.08.014

ing optimal control strategies. However, by ignoring the true dynamic nature of these processes, i.e., by using the classic, static MFA approach, important information on the flux dynamics is lost [13].

The methods in the second class do integrate this dynamic nature into the model structure by using the dMFA methodology and combining them with, e.g., EKF or MHE. Existing applications are focussed on estimating the true states, i.e., concentrations, from noisy data or estimating them if a measurement is not available. Goffaux et al. [14] compare the ensemble Kalman filter and an extended Kalman filter to estimate concentrations in a small-scale reconstruction of the energy metabolism of cells. Kawohl et al. [15] describe different methodologies (EKF, constrained EKF and MHE) to estimate states in a compartment model for *Streptomyces tendae*, combined with subsequent model predictive control based on these state estimates.

Based on these observations, there is a need for a true dynamic methodology in which, apart from the estimation of the states, also an accurate estimate of the fluxes in the metabolic reaction network can be found. The contribution of this work is the development of a methodology that combines a dMFA model with black-box expressions for the free fluxes, and a moving horizon estimation strategy that estimates both the states and the parameters in these black-box expressions. This way, a hybrid model structure, mechanistic on the stoichiometric level and black-box on the kinetic level, is developed which can be used for basically any organism for which a metabolic reaction network is available, and for a wide range of process conditions due

to the adaptive nature of the kinetic expressions, implemented through possibly time-variant parameters.

Before moving to the description of this methodology, an important remark has to be made. All methods described in this section, and of course also the method developed in this paper, rely on the availability of online measurements of extracellular concentrations. Although not yet mainstream in many bioprocesses, in recent years large progress has been made on this process analytical side. Online bioprocess monitoring is now possible using techniques like Raman spectroscopy [16], infrared spectroscopy [12], in situ techniques [17] and flow injection analysis [18]. In the following sections, the methods are developed based on the assumption that the necessary concentration measurements can be performed online.

This contribution is organized as follows. After this Introduction, the methodology is developed in the Material and Methods section. First, a short introduction to the dMFA model structure is given, after which the black-box flux expressions that are used in the dMFA model and that contain the parameters to be estimated, are specified. Finally, the specific MHE implementation used in this work is described. In the Results and Discussion section, the developed methodology is illustrated on two case studies: a small-scale case study, in which the different flux models are compared and a more detailed study on the effect of important model and method parameters is given, and a medium-scale network for *Escherichia coli*, in which the insights from the small-scale case study are used to estimate the fluxes in a more

realistic setting. Finally, the most important contributions of this paper are summarized in the Conclusions section.

2. Material and methods

2.1. The dMFA model structure

The general dMFA model structure is derived in [19] and [13]. It is based on a macroscopic description of the extracellular metabolites, along with a metabolic-reaction network-based model for the intracellular metabolites. These metabolic reaction networks describe the interactions between m_{ext} extracellular metabolites, m_{int} intracellular metabolites and the biomass, through n reactions, which are subdivided into n_{rev} reversible and n_{irr} irreversible reactions. All stoichiometric information in this network is summarized in the stoichiometric matrix $\mathbf{S}$, which is subdivided into the rows corresponding to intracellular metabolites $\mathbf{S}_{\text{int}}$, and the rows corresponding to extracellular metabolites and biomass $\mathbf{S}_{\text{e}}$. To also describe the irreversibilities in matrix form, an $(n_{\text{irr}} \times n)$ irreversibility matrix $\mathbf{I}_{\text{irr}}$ is set up, which selects the irreversible fluxes from the full set of fluxes.

The multi-scale model is simplified using the pseudo steady-state assumption on the intracellular level, as the dynamics at that level are much faster than the extracellular dynamics. This results in the following dynamic-algebraic system:

$$\frac{d\mathbf{x}(t)}{dt} = \mathbf{S}_{\text{e}} \cdot \hat{\mathbf{v}}(t) \cdot c_{\text{bio}}(t) \tag{1}$$

$$0 = \mathbf{S}_{\text{int}} \cdot \hat{\mathbf{v}}(t) \tag{2}$$

Postprint version of paper published in Journal of Process Control 2016, vol. 37, pages 1-20.
The content is identical to the published paper, but without the final typesetting by the publisher.

In this model $\mathbf{x}(t)$ is the $(m_{\text{ext}}+1\times 1)$ vector of concentration states for both the m_{ext} extracellular metabolites and the biomass, i.e.:

$$\mathbf{x}(t) = \begin{bmatrix} \mathbf{c}_{\text{ext}}(t) \\ c_{\text{bio}}(t) \end{bmatrix} \tag{3}$$

Furthermore, $\hat{\mathbf{v}}(t)$ is the $(n \times 1)$ vector of reaction rates, the so-called *fluxes*. In the majority of metabolic reaction networks, the number of intracellular metabolites is smaller than the number of reactions, making Equation (2) an underdetermined system of linear equations. The number of degrees of freedom d in the system equals the number of unknowns minus the number of independent equations, i.e., $d = n - rank(\mathbf{S}_{\text{int}})$. All solutions to this system can be written as a linear combination of a set of independent fluxes, called the *free fluxes*:

$$\hat{\mathbf{v}}(t) = \mathbf{K} \cdot \hat{\mathbf{u}}(t) \tag{4}$$

with $\mathbf{K}$ a suitable basis for the null space of $\mathbf{S}_{\text{int}}$ of dimensions $(n \times d)$ and $\hat{\mathbf{u}}(t)$ the $(d \times 1)$ vector of free fluxes. By substituting Equation (4) into Equation (1), the dynamic-algebraic system is turned into a truly dynamic model, which is referred to as the dMFA model structure:

$$\frac{d\mathbf{x}(t)}{dt} = \mathbf{S}_{\text{e}} \cdot \mathbf{K} \cdot \hat{\mathbf{u}}(t) \cdot \mathbf{q}_{\text{bio}}^{\intercal} \cdot \mathbf{x}(t) \tag{5}$$

Here, $\mathbf{q}_{\text{bio}}^{\intercal}$ selects the biomass from the full vector of state variables, i.e., $\mathbf{q}_{\text{bio}}^{\intercal} = \begin{bmatrix} 0 & 0 & \dots & 0 & 1 \end{bmatrix}$.

Apart from the system equations, also a number of constraints needs to be taken into account. As some of the reactions are irreversible, the fluxes

Journal homepage: http://www.journals.elsevier.com/journal-of-process-control
Original file available at: doi:10.1016/j.jprocont.2015.08.014

for these reactions should only be allowed to be positive, and this over the full time period under consideration. This is mathematically represented by adding extra algebraic states $\mathbf{z}(t)$, corresponding to the irreversible fluxes, to the problem and constraining these states to be positive:

$$\mathbf{z}(t) - \mathbf{I}_{\text{irr}} \cdot \mathbf{K} \cdot \hat{\mathbf{u}}(t) = 0 \tag{6}$$

$$\mathbf{z}(t) \geq 0 \tag{7}$$

Also, as extensively elaborated upon in [13], it is sometimes possible and beneficial to optimize the null space basis $\mathbf{K}$ instead of choosing it fixed a priori. In that case, all values in the $\mathbf{K}$ matrix are added as optimization variables, and the following constraints are added to make sure that the optimized $\mathbf{K}$ matrix is a basis for the null space of $\mathbf{S}_{\text{int}}$, and that the basis vectors are orthogonal:

$$\mathbf{S}_{\text{int}} \cdot \mathbf{K} = 0 \tag{8}$$

$$\mathbf{K}^{\intercal} \cdot \mathbf{K} - \mathbf{I} = 0 \tag{9}$$

2.2. Black-box flux models

The estimation model that is used for online metabolic flux estimation is the following:

$$\frac{d\mathbf{x}(t)}{dt} = \mathbf{S}_{\text{e}} \cdot \mathbf{K} \cdot \hat{\mathbf{u}}(\mathbf{x}, \mathbf{p}) \cdot \mathbf{q}_{bio}^{\intercal} \cdot \mathbf{x} + \boldsymbol{\tau}(\mathbf{x}, \mathbf{p}) \tag{10}$$

with $\boldsymbol{\tau}(\mathbf{x}, \mathbf{p})$ the transport terms which are dependent on the type of bioprocess under study. When a metabolic reaction network and the appropriate transport terms are determined, only the kinetic expressions for the free fluxes $\hat{\mathbf{u}}(\mathbf{x}, \mathbf{p})$ need to be determined. These can either be based on mechanistic knowledge about the intracellular reaction kinetics, and are in that case

highly specific for the organism at hand, or determined as black-box expressions, based on the online gathered data. In this work, the second option is chosen, as this results in an easy to implement, integrated way of flux estimation, as no specific flux expressions need to be identified. This methodology is also highly transferable between microorganisms, if a metabolic reaction network for these microorganisms is known.

Two black-box expressions for the flux model are used and compared in this paper: a linear function of the concentrations, and a nonlinear model that can describe both saturation and inhibition by the extracellular metabolites [20].

In the linear model, the free fluxes are described as linear functions of the extracellular concentrations:

$$\hat{\mathbf{u}}(t) = \mathbf{P}_{\mathrm{u}} \cdot \mathbf{x}_{\mathrm{ext}}(t) \tag{11}$$

with $\mathbf{P}_{\mathrm{u}}$ the $(d \times m_{\mathrm{ext}})$ matrix of linear coefficients, and $\mathbf{x}_{\mathrm{ext}}$ the $(m_{\mathrm{ext}} \times 1)$ vector of extracellular metabolite concentrations, i.e., without biomass. It is thus assumed that the specific fluxes are not directly depending on the biomass concentration anymore, as this is already taken care of by the multiplication with the biomass concentration in the dMFA model. In the end, the values in $\mathbf{P}_{\mathrm{u}}$ are the parameters that need to be estimated during the MHE procedure. There are no bounds on these parameters. This model will be referenced to as the *Linear* model.

The nonlinear model was developed by Haag et al. [20] as a general non-

Postprint version of paper published in Journal of Process Control 2016, vol. 37, pages 1-20.
The content is identical to the published paper, but without the final typesetting by the publisher.

linear kinetic model structure for bioreactions. It is specifically designed to capture the most important bioreaction phenomena, i.e., limitation, activation and inhibition, with as few parameters as possible:

$$\hat{u}_i(t) = u_{\max,i} \cdot \prod_{j=1}^{m_{\text{ext}}} \alpha_{ij}(c_{\text{ext},j}(t)) \tag{12}$$

with

$$\alpha_{ij}(c_{\text{ext},j}(t)) = \begin{cases} \frac{c_{\text{ext},j}}{c_{\text{ext},j}+K_{\text{H},ij}^2} & \text{if } K_{\text{H},ij} \geq 0 \\ \frac{1}{1+c_{\text{ext},j}\cdot K_{\text{H},ij}^2} & \text{if } K_{\text{H},ij} < 0 \end{cases} \tag{13}$$

In these equations, $\mathbf{u}_{\max}$ is the $(d \times 1)$ vector of maximum rate parameters, and $\mathbf{K}_{\text{H}}$ is the $(d \times m_{\text{ext}})$ matrix of Haag modulation parameters. If these parameters are positive, they describe a positive activation/saturation effect of the concentrations on the fluxes, if they are negative, an inhibition effect is represented. These two entities contain the set of parameters to be estimated by the MHE procedure. This model will be designated as the *Haag* model.

Although a predictive flux model structure, i.e., depending on the metabolite concentrations, is necessary to be able to use the dMFA model for predictive purposes, e.g., model predictive control, this is not necessary if one is only interested in the flux estimates. For this reason, a third flux model structure is studied to check if the dependency on the extracellular metabolite concentrations is really necessary to obtain an accurate estimate. In this nonpredictive model structure, the fluxes are just described as a constant over the estimation horizon:

$$\hat{\mathbf{u}}(t) = \mathbf{p}_{\text{u}} \tag{14}$$

Journal homepage: http://www.journals.elsevier.com/journal-of-process-control
Original file available at: doi:10.1016/j.jprocont.2015.08.014

with $\mathbf{p}_u$ the constant flux values, which are the parameters to be estimated by the MHE. This final model will be called the *NoFx* model, as in this case there is no dependency on $\mathbf{x}$.

2.3. Moving horizon estimation

Moving horizon estimation (MHE) is an optimization-based, online estimation technique for states and parameters [21]. At one point in time t_{L+N}, the states and parameters are estimated based on a subset of measurements ending with and including the measurement at t_{L+N}. This subset of $N+1$ measurements is called the *estimation horizon*. When a new measurement comes in, i.e., at time t_{L+N+1}, the new measurement is added to the horizon and the first measurement of the previous horizon is discarded, effectively moving the horizon one time step. This is shown graphically in Figure 1. MHE actually is an approximation of the so-called *full information estimation* problem, which takes into account all previous data up to t_{L+N}. The influence of this previous data, which is discarded in the MHE problem, is represented by the so-called *arrival cost* that is added to the objective function of the estimation problem.

2.3.1. Mathematical formulation

During the MHE procedure, at every time step a dynamic optimization problem has to be solved. The MHE formulation for this optimization problem used in this work is based on [22]. It features a combined state-parameter vector, process noise terms for both states and parameters, and an output noise term for the measurements. The formulation is mathematically repre-

sented as follows, at timepoint t_{L+N}:

$$\underset{\mathbf{x}_L^{\mathrm{c}},\mathbf{w}_L,\ldots,\mathbf{w}_{L+N-1}}{\text{minimize}} \sum_{j=L}^{L+N} \|\mathbf{m}_j - \mathbf{y}(t_j)\|_{\mathbf{V}}^2 + \sum_{j=L}^{L+N-1} \|\mathbf{w}_j\|_{\mathbf{W}}^2 + \|\mathbf{x}_L^{\mathrm{c}} - \bar{\mathbf{x}}_L^{\mathrm{c}}\|_{\bar{\mathbf{P}}_L}^2 \tag{15}$$

subject to:

$$\dot{\mathbf{x}}^{\mathrm{c}}(t) = \begin{bmatrix} \mathbf{f}(\mathbf{x}^{\mathrm{c}}) \\ 0 \end{bmatrix} + \begin{bmatrix} \boldsymbol{\omega}^{\mathrm{x}}(t) \\ \boldsymbol{\omega}^{\mathrm{p}}(t) \end{bmatrix} \tag{16}$$

$$\mathbf{x}^{\mathrm{c}}(0) = \mathbf{x}_L^{\mathrm{c}} \tag{17}$$

$$\mathbf{y}(t) = \mathbf{g}(\mathbf{x}^{\mathrm{c}}) \tag{18}$$

$$0 \leq \mathbf{h}(\mathbf{x}^{\mathrm{c}}) \tag{19}$$

As in [22], the norm in these equations is defined as follows:

$$\|\mathbf{a}\|_{\mathbf{A}}^2 = \mathbf{a}^{\intercal} \cdot \mathbf{A}^{\intercal} \cdot \mathbf{A} \cdot \mathbf{a} \tag{20}$$

In this problem, $\mathbf{x}^{\mathrm{c}}$ is the $(n_{\mathrm{x}} + n_{\mathrm{p}} \times 1)$ combined state and parameter vector, i.e.:

$$\mathbf{x}^{\mathrm{c}} = \begin{bmatrix} \mathbf{x} \\ \mathbf{p} \end{bmatrix} \tag{21}$$

$\mathbf{x}_L^{\mathrm{c}}$ is the value of this vector at time t_L, i.e., at the beginning of the horizon, and $\mathbf{w}_j$ is the $(n_{\mathrm{x}} + n_{\mathrm{p}} \times 1)$ combined state and parameter process noise vector at time t_j, i.e.:

$$\mathbf{w}_j = \begin{bmatrix} \mathbf{w}_j^{\mathrm{x}} \\ \mathbf{w}_j^{\mathrm{p}} \end{bmatrix} \tag{22}$$

Journal homepage: http://www.journals.elsevier.com/journal-of-process-control
Original file available at: doi:10.1016/j.jprocont.2015.08.014

Postprint version of paper published in Journal of Process Control 2016, vol. 37, pages 1-20.
The content is identical to the published paper, but without the final typesetting by the publisher.

There are four sets of constraints in this problem. The first set (Equation (16)) represents the *dynamic model*. In this equation, $\mathbf{f}(\mathbf{x}^{\mathrm{c}})$ is the $(n_{\mathrm{x}} \times 1)$ right hand side function for the states. For the parameters, these right hand sides are of course zero. To allow for mismatch between the real process and the estimation model used in MHE, and for (small) variations over time in the parameters, process noise terms are added to these right hand sides. The process noise is represented as an $(n_{\mathrm{x}} + n_{\mathrm{p}} \times 1)$ piecewise constant function $\boldsymbol{\omega}(t)$ with:

$$\boldsymbol{\omega}(t) = \sum_{j=L}^{L+N-1} \mathbf{w}_j \cdot \psi_j(t) \tag{23}$$

$$\psi_j(t) = \begin{cases} 1 & \text{if } t_j \leq t < t_{j+1} \\ 0 & \text{otherwise} \end{cases} \tag{24}$$

This function is again split into parts for the states and the parameters:

$$\boldsymbol{\omega}(t) = \begin{bmatrix} \boldsymbol{\omega}^{\mathrm{x}}(t) \\ \boldsymbol{\omega}^{\mathrm{p}}(t) \end{bmatrix} \tag{25}$$

The dynamic model constraints of course need to be discretized in some way to be able to solve the dynamic optimization problem. The second set of constraints stands for the *initial condition constraints* for the dynamic model (Equation (17)). The *output function* $\mathbf{y}(t)$ is defined in the third set of constraints (Equation (18)), with $\mathbf{g}(\mathbf{x}^{\mathrm{c}})$ any nonlinear $(n_{\mathrm{y}} \times 1)$ function of the states and parameters. Finally, also *general (non)linear constraints* (Equation (19)), like, e.g., the irreversibility constraints in the dMFA model, can be added to the constraint list.

Journal homepage: http://www.journals.elsevier.com/journal-of-process-control
Original file available at: doi:10.1016/j.jprocont.2015.08.014

The *first term* in the objective function minimizes the error between the measurements and the simulated output: $\mathbf{y}(t_j)$ is the $(n_\mathrm{y} \times 1)$ vector of model outputs at time t_j, $\mathbf{m}_j$ is the $(n_\mathrm{y} \times 1)$ vector with measurements at that time. This sum of squares is weighted with the $(n_\mathrm{y} \times n_\mathrm{y})$ matrix $\mathbf{V}$, which contains the inverses of the standard deviations of the measurements of the different outputs on its diagonal:

$$\mathbf{V} = \mathbf{\Sigma}_\mathrm{V}^{-1} = \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 & 0 \\ 0 & \sigma_2 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \sigma_{n_\mathrm{y}-1} & 0 \\ 0 & 0 & \cdots & 0 & \sigma_{n_\mathrm{y}} \end{bmatrix}^{-1} \tag{26}$$

with σ_i the measurement standard deviation corresponding to output i. In principle, these standard deviations can be varying over time, but for this work, they are kept constant.

The *second term* in the objective function minimizes the estimated process noise. This term is weighted with the $(n_\mathrm{x} + n_\mathrm{p} \times n_\mathrm{x} + n_\mathrm{p})$ matrix $\mathbf{W}$, with:

$$\mathbf{W} = \begin{bmatrix} \mathbf{W}_\mathrm{x} & 0 \\ 0 & \mathbf{W}_\mathrm{p} \end{bmatrix} \tag{27}$$

with $\mathbf{W}_\mathrm{x}$ the $(n_\mathrm{x} \times n_\mathrm{x})$ matrix with the inverses of the standard deviations of the process noise for the states on the diagonal, and $\mathbf{W}_\mathrm{p}$ the $(n_\mathrm{p} \times n_\mathrm{p})$ matrix with the inverses of the standard deviations of the process noise for the parameters on the diagonal. Both are defined in the same way as $\mathbf{V}$ in

Equation (26), i.e.:

$$\mathbf{W} = \mathbf{\Sigma}_{\mathrm{W}}^{-1} \tag{28}$$

Also these standard deviations are kept constant over time. All noise terms is this formulation are assumed to be normally distributed, independent, additive noise.

The *last term* in the objective function is the arrival cost. In this term, $\bar{\mathbf{x}}_L^{\mathrm{c}}$ is the $(n_{\mathrm{x}} + n_{\mathrm{p}} \times 1)$ estimated combined state-parameter vector at time t_L. This last term is weighted with the $(n_{\mathrm{x}} + n_{\mathrm{p}} \times n_{\mathrm{x}} + n_{\mathrm{p}})$ matrix $\bar{\mathbf{P}}_L$, which represents the inverse square root of the estimated (co)variance of the estimated combined state-parameter vector. This matrix will be updated in each iteration of the MHE procedure, along with $\bar{\mathbf{x}}^{\mathrm{c}}$. To start the MHE procedure, initial guesses are needed for both these MHE parameters, as of course no estimate is yet available for them at that point. These initial guesses are indicated as $\bar{\mathbf{x}}_0^{\mathrm{c}}$ and $\bar{\mathbf{P}}_0$.

2.3.2. Arrival cost approximation

Different methods for updating the arrival cost parameters, i.e., $\bar{\mathbf{x}}^{\mathrm{c}}$ and $\bar{\mathbf{P}}$, are described in literature. The conventional method is to use an extended Kalman filter to propagate the a priori state estimate and covariance [23]. In the case of a nonlinear system, or when bounds are present, however, errors are introduced in the arrival cost term, which makes the use of longer horizon lengths necessary [24]. This problem can be overcome by using sampling-based filters for the arrival cost update, like, e.g., the unscented Kalman filter of particle filter [25, 26].

Postprint version of paper published in Journal of Process Control 2016, vol. 37, pages 1-20.
The content is identical to the published paper, but without the final typesetting by the publisher.

The method which is used in this work is described in [22]. As opposed to the filter based updates, it is motivated slightly different, but results in an easy to implement and efficient update formulation. Here, it is adopted to a continuous system instead of a discrete system. It is based on the ideal nonlinear arrival cost C:

$$C(\mathbf{x}^{\mathrm{c}}_{L+1}) = \min_{\mathbf{x}^{\mathrm{c}}_{L}} \|\mathbf{m}_L - \mathbf{y}(t_L)\|^2_{\mathbf{V}} + \left\| \begin{matrix} \mathbf{w}^{\mathrm{x}}_L \\ \mathbf{w}^{\mathrm{p}}_L \end{matrix} \right\|^2_{\mathbf{W}} + \|\mathbf{x}^{\mathrm{c}}_L - \bar{\mathbf{x}}^{\mathrm{c}}_L\|^2_{\bar{\mathbf{P}}_L} \tag{29}$$

subject to:

$$\mathbf{w}^{\mathrm{x}}_L = \dot{\mathbf{x}}_L - \mathbf{f}(\mathbf{x}^{\mathrm{c}}_L) \tag{30}$$

$$\mathbf{w}^{\mathrm{p}}_L = \dot{\mathbf{p}}_L \tag{31}$$

This nonlinear arrival cost is now approximated by a quadratic term, which contains all information in the interval $[t_L, t_{L+1}]$ and can be used to set up the MHE problem for the next measurement point. To do this, the nonlinear functions and derivatives in Equations (29)-(31) are linearized as follows:

$$\mathbf{f}(\mathbf{x}^{\mathrm{c}}) \approx \underbrace{\mathbf{f}(\mathbf{x}^{\mathrm{c}*}) - \frac{d\mathbf{f}(\mathbf{x}^{\mathrm{c}*})}{d\mathbf{x}^{\mathrm{c}}} \cdot \mathbf{x}^{\mathrm{c}*}}_{\tilde{\mathbf{f}}} + \frac{d\mathbf{f}(\mathbf{x}^{\mathrm{c}*})}{d\mathbf{x}^{\mathrm{c}}} \cdot \mathbf{x}^{\mathrm{c}} \tag{32}$$

$$\mathbf{y}(\mathbf{x}^{\mathrm{c}}) \approx \underbrace{\mathbf{y}(\mathbf{x}^{\mathrm{c}*}) - \frac{d\mathbf{y}(\mathbf{x}^{\mathrm{c}*})}{d\mathbf{x}^{\mathrm{c}}} \cdot \mathbf{x}^{\mathrm{c}*}}_{\tilde{\mathbf{y}}} + \frac{d\mathbf{y}(\mathbf{x}^{\mathrm{c}*})}{d\mathbf{x}^{\mathrm{c}}} \cdot \mathbf{x}^{\mathrm{c}} \tag{33}$$

$$\dot{\mathbf{x}} \approx \frac{\mathbf{x}_{L+1} - \mathbf{x}_L}{\Delta t} \tag{34}$$

$$\dot{\mathbf{p}} \approx \frac{\mathbf{p}_{L+1} - \mathbf{p}_L}{\Delta t} \tag{35}$$

with Δt the time between two consecutive measurements, $\mathbf{x}^{\mathrm{c}*}$ the best available estimate. By substituting these linearizations into the ideal arrival cost

Journal homepage: http://www.journals.elsevier.com/journal-of-process-control
Original file available at: doi:10.1016/j.jprocont.2015.08.014

Postprint version of paper published in Journal of Process Control 2016, vol. 37, pages 1-20.
The content is identical to the published paper, but without the final typesetting by the publisher.

(Equation (29)), the following linear least squares problem arises:

$$\min_{\mathbf{x}_L^c} \left|\left| \mathbf{A} \cdot \begin{bmatrix} \mathbf{x}_L^c \\ \mathbf{x}_{L+1}^c \end{bmatrix} - \mathbf{b} \right|\right|_2^2 \tag{36}$$

with:

$$\mathbf{A} = \begin{bmatrix} -\mathbf{V} \cdot \frac{d\mathbf{y}(\mathbf{x}^{c*})}{d\mathbf{x}^c} & 0 \\ -\mathbf{W} \cdot \left(\begin{bmatrix} \frac{d\mathbf{f}(\mathbf{x}^{c*})}{d\mathbf{x}^c} \\ 0 \end{bmatrix} + \frac{1}{\Delta t} \cdot \mathbf{I} \right) & \frac{\mathbf{W}}{\Delta t} \\ \bar{\mathbf{P}}_L & 0 \end{bmatrix} \tag{37}$$

$$\mathbf{b} = \begin{bmatrix} \mathbf{V} \cdot (\tilde{\mathbf{y}} - \mathbf{m}_L) \\ \mathbf{W} \cdot \tilde{\mathbf{f}} \\ \bar{\mathbf{P}}_L \cdot \bar{\mathbf{x}}_L^c \end{bmatrix} \tag{38}$$

Using the QR decomposition of the matrix A:

$$\mathbf{A} = \begin{bmatrix} \mathbf{Q}_1 & \mathbf{Q}_2 & \mathbf{Q}_3 \end{bmatrix} \cdot \begin{bmatrix} \mathbf{R}_1 & \mathbf{R}_{12} \\ 0 & \mathbf{R}_2 \\ 0 & 0 \end{bmatrix} \tag{39}$$

this problem has the following analytic solution:

$$C'(\mathbf{x}_{L+1}^c) = \|\mathbf{Q}_3 \cdot \mathbf{b}\|_2^2 + \|\mathbf{Q}_2 \cdot \mathbf{b} + \mathbf{R}_2 \cdot \mathbf{x}_{L+1}^c\|_2^2 \tag{40}$$

As the first term is a constant, it does not influence the approximated arrival cost. The second term can be transformed into the wanted quadratic form for the arrival cost:

$$C'(\mathbf{x}_{L+1}^c) = \|\bar{\mathbf{P}}_{L+1} \cdot (\mathbf{x}_{L+1}^c - \bar{\mathbf{x}}_{L+1}^c)\|_2^2 \tag{41}$$

Journal homepage: http://www.journals.elsevier.com/journal-of-process-control
Original file available at: doi:10.1016/j.jprocont.2015.08.014

with:

$$\bar{\mathbf{P}}_{L+1} = \mathbf{R}_2 \tag{42}$$

$$\bar{\mathbf{x}}^{\mathrm{c}}_{L+1} = \mathbf{R}_2^{-1} \cdot \mathbf{Q}_2 \cdot \mathbf{b} \tag{43}$$

These updated arrival cost parameters are then used to set up the MHE objective function (Equation (15)) for the next sampling instant.

2.3.3. Numerical implementation

As in [13] and [27], all nonlinear dynamic optimization problems are solved using direct collocation on finite elements. Cubic Lagrange polynomials were utilized, with finite element borders on the sampling time points, i.e., for an horizon length of 20, 20 finite elements were used. The resulting nonlinear programming problems are solved using the interior-point optimization routine IPOPT [28]. Gradient, Jacobians and Hessian are generated exactly using automatic differentiation with CasADi [29].

3. Results and discussion

3.1. Small-scale case study

3.1.1. Description of case study

The small-scale case study uses the same network as the small-scale case study in [13], and is shown in Figure 2, along with the corresponding stoichiometric, irreversibility and null space basis matrices. This network consists of 3 extracellular metabolites and biomass, 4 intracellular metabolites and 7 fluxes. Thus, the number of free fluxes is 3. For the simulation of the

measurements, these were chosen as fluxes 1, 4 and 5, as in [13]. The system under study here is a continuous bioreactor to which a mixture of two nutrients, A and E, is fed, from two tanks that contain solutions of A and E with concentrations $c_{\text{in,A}}$ and $c_{\text{in,E}}$. The specific composition of the input flow can be controlled. This mix is represented by two control variables r_A and r_E, which correspond to the fraction of the input flow that is taken from the corresponding nutrient tank. For that reason, the sum of the two controls can be at most one, and if it is lower than one, the remainder of the mix is made up of nutrient-free medium. This way, the flow rates in and out of the reactor are always balanced and it is not necessary to also set up a volume balance.

Two different models are used in this case study: a *simulation model*, which corresponds to the *real* process and which is used to simulate the measurements, and an *estimation model*, which is an approximation of the real model since the real flux expressions are not known, and which is used to estimate the fluxes.

The simulation model is the following:

$$\frac{d\mathbf{x}(t)}{dt} = \mathbf{S}_{\text{e}} \cdot \mathbf{K} \cdot \mathbf{u}_{\text{sim}}(t) \cdot \mathbf{q}_{\text{bio}}^{\intercal} \cdot \mathbf{x}(t) + (\mathbf{X}_{\text{in}} \cdot \mathbf{r}(t) - \mathbf{x}(t)) \cdot D \qquad (44)$$

with $\mathbf{X}_{\text{in}}$ the (4×2) matrix of inlet tank concentrations for the different metabolites [mol/L], $\mathbf{r}$ the (2×1) vector of control variables manipulating the feed composition that is sent to the tank, and D the dilution rate [1/h].

The dilution rate is fixed to 0.1 h^{-1}. For this case, the following holds:

$$\mathbf{X}_{\text{in}} = \begin{bmatrix} c_{\text{in,A}} & 0 \\ 0 & c_{\text{in,E}} \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \tag{45}$$

$$\mathbf{r}(t) = \begin{bmatrix} r_A(t) \\ r_E(t) \end{bmatrix} \tag{46}$$

The simulation model flux expressions $\mathbf{u}_{\text{sim}}(t)$ are chosen as:

$$u_{1,\text{sim}} = \frac{c_{\text{Aext}}}{1.5 + c_{\text{Aext}}} \tag{47}$$

$$u_{4,\text{sim}} = 0.2 \cdot \frac{c_{\text{Eext}}}{3 + c_{\text{Eext}}} \tag{48}$$

$$u_{5,\text{sim}} = \frac{1}{1 + c_{\text{Fext}}} \tag{49}$$

Using the simulation model, measurements were generated for a time range of 160 hours, with measurements of all states every 6 minutes. Input profiles for the controls were chosen as depicted in Figure 3. Additive, indepent, normal measurement noise was added to the outputs of the true model, and also additive, independent, normal process noise was added to the right hand sides of the ODE system. An overview of the states, along with their initial condition, process and measurement noises is given in Table 1. The final set of measurements is shown in Figure 4.

For the estimation model, all features of the simulation model were assumed to be known, except for the unknown flux dynamics. These are described by the black-box approximations described in Section 2.2. As a

benchmark, estimation with the true flux kinetics model structure is also considered. For this case, the true flux model (Equations (47) to (49)) is parameterized as follows:

$$\hat{u}_1 = u_{\max,1} \cdot \frac{c_{\text{Aext}}}{K_{\text{M},1} + c_{\text{Aext}}} \tag{50}$$

$$\hat{u}_4 = u_{\max,4} \cdot \frac{c_{\text{Eext}}}{K_{\text{M},4} + c_{\text{Eext}}} \tag{51}$$

$$\hat{u}_5 = u_{\max,5} \cdot \frac{1}{K_{\text{M},5} + c_{\text{Fext}}} \tag{52}$$

In total, 7 scenarios are considered: estimation with the true model, the *NoFx* model with fixed and free (optimized) $\mathbf{K}$ matrix, the *linear* model with fixed and free $\mathbf{K}$ matrix, and the *Haag* model with fixed and free $\mathbf{K}$ matrix. In the cases with fixed $\mathbf{K}$ matrix, this is chosen as the rational basis according to free fluxes 1, 4 and 5. An overview of these scenarios is given in Table 2.

3.1.2. MHE problem and settings

The MHE problem for this case study at a point in time t_{L+N} is the following:

$$\underset{\mathbf{x}_L^{\text{c}}, \mathbf{w}_L, \ldots, \mathbf{w}_{L+N-1}}{\text{minimize}} \sum_{j=L}^{L+N} \|\mathbf{m}_j - \mathbf{y}(t_j)\|_{\mathbf{V}}^2 + \sum_{j=L}^{L+N-1} \|\mathbf{w}_j\|_{\mathbf{W}}^2 + \|\mathbf{x}_L^{\text{c}} - \bar{\mathbf{x}}_L^{\text{c}}\|_{\bar{\mathbf{P}}_L}^2 \tag{53}$$

Journal homepage: http://www.journals.elsevier.com/journal-of-process-control
Original file available at: doi:10.1016/j.jprocont.2015.08.014

subject to:

$$\dot{\mathbf{x}}(t) = \mathbf{S}_\mathrm{e} \cdot \mathbf{K} \cdot \hat{\mathbf{u}}(\mathbf{x}, \mathbf{p}) \cdot \mathbf{q}_\mathrm{bio}^\mathsf{T} \cdot \mathbf{x}(t) + (\mathbf{X}_\mathrm{in} \cdot \mathbf{r} - \mathbf{x}(t)) \cdot D + \boldsymbol{\omega}_\mathrm{x}(t) \tag{54}$$

$$\dot{\mathbf{p}}(t) = \boldsymbol{\omega}_\mathrm{p}(t) \tag{55}$$

$$\mathbf{x}(0) = \mathbf{x}_L \tag{56}$$

$$\mathbf{p}(0) = \mathbf{p}_L \tag{57}$$

$$0 = \mathbf{z}(t) - \mathbf{I}_\mathrm{irr} \cdot \mathbf{K} \cdot \hat{\mathbf{u}}(\mathbf{x}, \mathbf{p}) \tag{58}$$

$$\mathbf{y}(t) = \mathbf{x}(t) \tag{59}$$

$$\mathbf{x}(t) \geq 0, \mathbf{z}(t) \geq 0 \tag{60}$$

In the scenarios with a free $\mathbf{K}$ matrix, the elements of $\mathbf{K}$ are added as degrees of freedom, while the null space and orthogonality constraints (Equations (8) and (9)) are added as constraints:

$$\mathbf{S}_\mathrm{int} \cdot \mathbf{K} = 0 \tag{61}$$

$$\mathbf{K}^T \cdot \mathbf{K} - \mathbf{I} = 0 \tag{62}$$

For all cases, the MHE horizon was chosen as 20, i.e., 20 measurement intervals and thus 21 measurement points in the horizon. An analysis of the horizon length revealed the influence of the horizon length on the estimation accuracy to be insignificant when compared to the influence of the parameter noise on the estimates. This analysis was carried out for horizon lengths between 10 and 50. For this reason, the horizon length was chosen arbitrarily. The $\mathbf{V}$ and $\mathbf{W}_\mathrm{x}$ matrices are set up with the noise standard deviations as defined in Table 1. The choice of the estimated noise level for the parameters is described in the next section. Lastly, initial guesses need to be made for the arrival cost parameters $\bar{\mathbf{x}}_0^\mathrm{c}$ and $\bar{\mathbf{P}}_0$. The state estimate is initialized at

Journal homepage: http://www.journals.elsevier.com/journal-of-process-control
Original file available at: doi:10.1016/j.jprocont.2015.08.014

the measurement at $t = 0$, as all extracellular metabolites and biomass are measurements, and the parameter estimates are dependent on the scenario and are given in Table 2. Finally, the standard deviations to be used in $\bar{\mathbf{P}}_0$ are chosen equal to the measurement standard deviation for the states, and are chosen as 1.0 for the parameters.

3.1.3. Optimal parameter noise selection

As said before, the choice of the parameter noise standard deviations to be used in the weighting matrix $\mathbf{W}_\mathrm{p}$ is not trivial, and can, intuitively, have a large influence on the final estimates. If they are chosen low, the parameters will not be able to change a lot over the course of the experiment, while if high standard deviations are chosen, parameters can vary very quickly, possibly introducing noise in the estimates as well. To empirically assess the influence of this MHE parameter on the estimates, the MHE procedure was executed for different levels of standard deviation for the estimated parameter noise ranging between 10^{-1} and 10^{-5}. In all cases, the noise level was the same for all parameters in the model. Estimation quality is assessed by calculating the mean squared errors (MSE) between simulated and estimated fluxes and states.

3.1.4. Results for flux estimation

First, the results regarding the estimation of the fluxes are depicted. It is important to note that estimates of both the fluxes and states are gathered at the same time, i.e., these are not separate estimation problems, but they are discussed here separately as they can serve a different purpose, flux estimation being more important for monitoring purposes and state estimation

being the basis for model-based predictive control.

The results for the different scenarios and different parameter noise levels are shown in Figure 5. The best results for each case are also summarized in Table 3. From these results, the following observations can be made.

The *Real* scenario is the reference case, as it expresses the true model structure. Since the parameters in this model represent the same, constant parameters used in the simulation, it is logical that the estimates are very accurate, and that the estimates get better when lowering the parameter noise, as the true values of the parameters are constant.

The *NoFx* case is on the other side of the spectrum. In this case, the parameters are just the final free fluxes to be estimated, and as these vary significantly in time, one can expect that a large amount of parameter noise is necessary. From the figure it is clear that there is an optimum parameter noise value where the error between simulated (the *true* process) and estimated fluxes is minimal.

In the *Linear* and *Haag* scenarios, some interesting results can be observed. The trend in the *Linear* case more resembles the trend in the *Real* case, i.e., decreasing towards low parameter noise, while in the *Haag* case, the opposite is true, i.e., the results resemble the *NoFx* case more. This is a bit counterintuitive as one expects the parameters to vary not so much in the *Haag* scenario, as the model structure is closer to the true model. The

Haag model, however, only gives a negligible improvement over not using state-dependent fluxes at all. The *Linear* model on the other hand, shows a very accurate estimation performance, coming close to estimation with the true flux model.

The *Linear* model also is more robust to a bad choice of the parameter noise, as long as the parameter noise is chosen small enough, as opposed to the *Haag* model, where on both sides of the optimum value, the results are very sensitive to deviations from this optimum parameter noise. If the parameter noise in the *Linear* model is chosen too big, however, the sensitivity is very high. This probably also happens in the *Haag* model if parameter noise values higher than 0.1 are chosen, but no simulations were performed in that region.

The trade-off that is underlying the choice of the parameter noise level, is most clearly shown in Figure 6. When the parameter noise is chosen too large, as on the left, there is also a large amount of noise on the estimated fluxes. If on the other hand the parameter noise is chosen too small, the parameters are not allowed to vary enough over the course of time and the estimate is far off. This effect is most clearly visible in the *Haag* case. The optimal choice provides an acceptable trade-off between variance and estimate quality.

In almost all points in these figures, the estimation with a free $\mathbf{K}$ matrix shows better or equal performance as estimation with the true $\mathbf{K}$ matrix.

This is an indication that the MHE procedure can be run perfectly without having to choose a basis for the null space of the stoichiometric matrix a priori. This is an important point as it makes the methodology self-contained and easily usable for different microorganisms or strains, without having to go into much detail regarding the metabolic reaction network and the chosen basis.

Finally, also in Figures 7 and 8, depicting the estimated fluxes over time for the different scenarios, these observations are clearly discernible. For both the fixed and free $\mathbf{K}$ matrices, the *Linear* model gets very close to the real flux profiles, after an initial period because of the initial guess for the estimated parameters and covariance estimate, while the *Haag* model captures the general trend, but gives estimates with more noise, which are barely better than in the *NoFx* scenario. These results are a first indication that the combination of a dMFA model structure with a linear flux model and moving horizon estimation is an excellent methodology for online estimation of metabolic fluxes for monitoring and control of bioprocesses.

3.1.5. Results for state estimation

The performance of the proposed methodology for state estimation, as a way to monitor unmeasured states as a software sensor, for data reconciliation or as an input for model-based predictive control, is also studied. Again, the mean squared errors between simulated and estimated states for the different scenarios and parameter noise levels are given in Figure 9 and Table 4.

Similar conclusions can be drawn from these figures as for the flux estimation. The biggest difference lies in the fact that in this case, also the *Linear* scenario exhibits a clear minimum, while in the *Haag* case, the results now yield a far flatter profile on the right of the optimum, i.e., in the larger parameter noise region. When looking at the state profiles at, and left and right of the optimum (Figure 10), though, the differences are minimal. For the *Linear* case, the estimate left of the optimum is slightly worse than at the optimum, but still acceptable, while for the *Haag* scenario, the estimates are closer to the true profiles, but a bit more noise is introduced. Also for state estimation, the *Linear* model is clearly the top contender, but the differences are far less pronounced, probably because all states are measured. The results for this model are also far less sensitive to changes in the parameter noise level than in the flux estimation. These findings also show in Figures 11 and 12, where the estimated profiles for the different scenarios and all fluxes are given.

Again, also when the $\mathbf{K}$ matrix is not fixed a priori, the estimates are accurate, but also here, the differences are not as pronounced as in the flux estimation.

3.1.6. Computational requirements

Apart from the estimation performance, also the computational performance has been assessed. These results are presented in Table 5. These times are achieved when running the algorithm on one core of an eight-core Intel i7-3770 CPU at 3.40 Ghz. Of course, although these numbers are for this small-scale case study, and are not attainable in real-life situations due

to bigger networks, more states, fluxes and parameters, the relative differences in computational performance between the different scenarios are still representative. The intuitive understanding that the *Linear* model must be easier to estimate than the nonlinear *Haag* model, also clearly shines through in these results. This is another reason to choose the *Linear* model, for the current methodology, over the *Haag* model, as it gives better results in less time. Finally, it can also be seen that the estimation of an optimal $\mathbf{K}$ matrix, as in the offline dMFA method, introduces a large computational burden. It is still to be checked whether the online determination of this null space basis is possible also for larger networks. There is, however, still quite some room on the computational part as bioprocesses are typically slow processes and thus sampling frequencies can be in the range of minutes or hours.

3.2. Realistic medium-scale network for E. coli

This methodology is also validated on a more realistic, medium-scale case study. The detailed analysis that was performed on the small-scale case study is not repeated, but based on the results obtained there, the *Linear* model is identified as a more suitable flux model for the current methodology. Thus, the estimations for this larger network are only performed with this *Linear* model as the black-box flux model, and with a parameter noise level that is chosen based on the results obtained for the small-scale case study. The implementation for a larger network with a more realistic measurement scenario is important to check the performance of the proposed methodology on both the estimation and the computational level.

Journal homepage: http://www.journals.elsevier.com/journal-of-process-control
Original file available at: doi:10.1016/j.jprocont.2015.08.014

3.2.1. Description of case study

The network used in the medium-scale case study is an adapted version of the core *E. coli* model [30] with the reactions for PDO production as in [31] included. The full set of reactions can be found in the Supplementary data. The resulting network contains 45 intracellular metabolites, 10 extracellular metabolites (including biomass), 50 fluxes and 6 free fluxes. The full set of state variables is shown in Table 6. Apart from the medium concentrations, also the headspace concentrations of oxygen and carbon dioxide were modeled. Again, a continuous bioreactor setup is assumed. The control variables are in this case the dilution rate and the inlet concentrations of fructose, glucose and glutamine. The feed medium also contains a fixed amount of phosphate, oxygen and carbon dioxide.

In this case study, the simulation model was chosen to be a dynamic flux balance analysis (dFBA) model [32]. This is a predictive model structure based on the pseudo steady-state assumption, the assumption that the cell tries to maximize its growth rate, and kinetic expression for the uptake fluxes. Mathematically, it is represented as a dynamic optimization problem with the maximization of the biomass flux as optimization objective and inequality constraints representing the maximal uptake rates of the different extracellular metabolites. Recently, this type of model has been applied in an industrial setting [33].

The choice for a dFBA model as the simulation model also introduces a mismatch between the simulation and estimation models. This mismatch is

typical of a realistic situation, as in real-life applications, the true model is never known, and there always is a mismatch between the model and the true process dynamics. The dFBA model that is used in this work is the following:

$$\underset{\mathbf{x}(t),\mathbf{v}(t)}{\text{maximize}}\ v_{\text{bio}} = v_0 \tag{63}$$

Journal homepage: http://www.journals.elsevier.com/journal-of-process-control
Original file available at: doi:10.1016/j.jprocont.2015.08.014

Postprint version of paper published in Journal of Process Control 2016, vol. 37, pages 1-20.
The content is identical to the published paper, but without the final typesetting by the publisher.

subject to:

$$\frac{d\mathbf{x}_{\text{meta}}(t)}{dt} = \mathbf{S}_{\text{e,meta}} \cdot \mathbf{v}(t) \cdot \mathbf{q}_{\text{bio}}^T \cdot \mathbf{x}_{\text{meta}}(t) + (\mathbf{x}_{\text{meta,in}} - \mathbf{x}_{\text{meta}}(t)) \cdot D(t) \tag{64}$$

$$\frac{d\mathbf{x}_{\text{diss}}(t)}{dt} = \mathbf{S}_{\text{e,diss}} \cdot \mathbf{v}(t) \cdot \mathbf{q}_{\text{bio}}^T \cdot \mathbf{x}_{\text{diss}}(t) + (\mathbf{x}_{\text{diss,in}} - \mathbf{x}_{\text{diss}}(t)) \cdot D(t) + \mathbf{K}_{\text{l}} \cdot (\mathbf{x}_{\text{diss}}^* - \mathbf{x}_{\text{diss}}) \tag{65}$$

$$\frac{d\mathbf{x}_{\text{head}}(t)}{dt} = -\frac{\mathbf{K}_{\text{l}} \cdot (\mathbf{x}_{\text{diss}}^* - \mathbf{x}_{\text{diss}}) \cdot V_{\text{liq}}}{M \cdot V_{\text{head}}} + \frac{\mathbf{x}_{\text{head,in}} \cdot F}{V_{\text{head}}} - \frac{\mathbf{x}_{\text{head}}(t) \cdot F}{V_{\text{head}}} \cdot \frac{1 - \sum_i \mathbf{x}_{\text{head,in},i}}{1 - \sum_i \mathbf{x}_{\text{head},i}(t)} \tag{66}$$

$$v_{46}(t) \leq 16.8 \cdot \frac{c_{\text{Phos}}}{10.0 + c_{\text{Phos}}} \cdot \frac{1}{3.0 + c_{\text{Ace}}} \tag{67}$$

$$v_{47}(t) \leq 2.0 \cdot \frac{c_{\text{Fru}}}{15.0 + c_{\text{Fru}}} \tag{68}$$

$$v_{48}(t) \leq 4.0 \cdot \frac{c_{\text{Glc}}}{7.6 + c_{\text{Glc}}} \tag{69}$$

$$v_{49}(t) \leq 3.0 \cdot \frac{c_{\text{Gln}}}{10.0 + c_{\text{Gln}}} \tag{70}$$

$$v_{50}(t) \leq 18.0 \cdot \frac{c_{\text{O}_2}}{0.14 + c_{\text{O}_2}} \tag{71}$$

$$0 \leq I_{\text{irr}} \cdot \mathbf{v}(t) \tag{72}$$

$$\mathbf{x}_{\text{meta}}(0) = \mathbf{x}_{\text{meta,0}} \tag{73}$$

$$\mathbf{x}_{\text{diss}}(0) = \mathbf{x}_{\text{diss,0}} \tag{74}$$

$$\mathbf{x}_{\text{head}}(0) = \mathbf{x}_{\text{head,0}} \tag{75}$$

with $\mathbf{x}_{\text{meta}}$, $\mathbf{x}_{\text{diss}}$ and $\mathbf{x}_{\text{head}}$ the (8×1) vector of medium metabolite concentra-

Journal homepage: http://www.journals.elsevier.com/journal-of-process-control
Original file available at: doi:10.1016/j.jprocont.2015.08.014

tions [mmol/L], the (2×1) vector of dissolved gas concentrations [mmol/L], and the (2×1) vector of headspace gas mole fractions [-], respectively, as indicated in Table 6. The corresponding symbols with an $_{\text{in}}$ subscript indicate the feed concentration vector for that set of metabolites, and the ones with a $_0$ subscript indicate the starting values for the simulation. $\mathbf{S}_{\text{e,meta}}$ and $\mathbf{S}_{\text{e,diss}}$ are the rows of $\mathbf{S}_{\text{e}}$ corresponding to the medium metabolites and the dissolved gases, respectively. V_{liq} and V_{head} are the volumes of medium and headspace in the reactor [L], respectively. M is the reciprocal of the ideal gas molar volume [mmol/L], and F is the inlet gas flow rate [L/h]. Finally, $\mathbf{x}^*_{\text{diss}}$ is the saturation concentration of the dissolved gases [mmol/h], and $\mathbf{K}_{\text{l}}$ is a (2×2) diagonal matrix with the $k_{\text{l}}a$ values for oxygen and carbon dioxide:

$$\mathbf{K}_{\text{l}} = \begin{bmatrix} (k_{\text{l}}a)_{\text{O}_2} & 0 \\ 0 & (k_{\text{l}}a)_{\text{CO}_2} \end{bmatrix} \tag{76}$$

The numeric values for the different simulation model parameters are given in Table 7.

It is important to note that, in this case study, no $\mathbf{K}$ matrix has been chosen a priori for the simulation model, as it involves a dynamic flux balance analysis model. This means that also in the estimation, there is no a priori best choice for the $\mathbf{K}$ matrix, as opposed to the previous case study where the simulation model was run with a predefined $\mathbf{K}$ matrix.

Based on this simulation model, measurements were generated for a time range of 300 hours, with measurements as indicated in Table 6 every 6 minutes. The chosen input profiles for the controls are given in Figure 13. Addi-

tive, independent, normal measurement noise was added to the outputs of the true model, and also additive, independent, normal process noise was added to the right hand sides of the ODE system. An overview of the states, along with initial conditions, process and measurement noises is given in Table 6. The final set of measurements is displayed in Figure 14.

3.2.2. MHE problem and settings

Based on the results for the small-scale case study, the estimation for this case study was only performed with the linear flux model. This model describes a linear relationship between the free fluxes and the states except biomass:

$$\hat{\mathbf{u}}(t) = \mathbf{P}_{\mathrm{u}} \cdot \mathbf{x}_{\mathrm{ext}}(t) \tag{77}$$

In this case, $\mathbf{x}_{\mathrm{ext}}(t)$ is a (9×1) vector consisting of all metabolite and dissolved concentration states except biomass, and thus $\mathbf{P}_{\mathrm{u}}$ is a (6×9) matrix with the linear coefficients, which are the parameters to estimate. Furthermore, also the $k_{\mathrm{l}}a$ values for both oxygen and carbon dioxide were estimated, as these are typically hard to determine experimentally, and only measurements of the headspace gases are available. This makes in total 56 parameters that have to be estimated. This results in the following MHE problem formulation at a point in time t_{L+N}:

$$\underset{\mathbf{x}_L^{\mathrm{c}},\mathbf{w}_L,\ldots,\mathbf{w}_{L+N-1}}{\text{minimize}} \sum_{j=L}^{L+N} \|\mathbf{m}_j - \mathbf{y}(t_j)\|_{\mathbf{V}}^2 + \sum_{j=L}^{L+N-1} \|\mathbf{w}_j\|_{\mathbf{W}}^2 + \|\mathbf{x}_L^{\mathrm{c}} - \bar{\mathbf{x}}_L^{\mathrm{c}}\|_{\bar{\mathbf{P}}_L}^2 \tag{78}$$

subject to:

$$\dot{\mathbf{x}}(t) = \begin{bmatrix} \dot{\mathbf{x}}_{\text{meta}}(t) \\ \dot{\mathbf{x}}_{\text{diss}}(t) \\ \dot{\mathbf{x}}_{\text{head}}(t) \end{bmatrix} = \begin{bmatrix} \mathbf{f}_{\text{meta}}(\mathbf{x}, \mathbf{p}) \\ \mathbf{f}_{\text{diss}}(\mathbf{x}, \mathbf{p}) \\ \mathbf{f}_{\text{head}}(\mathbf{x}, \mathbf{p}) \end{bmatrix} + \boldsymbol{\omega}_{\text{x}}(t) \tag{79}$$

$$\dot{\mathbf{p}}(t) = \boldsymbol{\omega}_{\text{p}}(t) \tag{80}$$

$$\mathbf{x}(0) = \mathbf{x}_L \tag{81}$$

$$\mathbf{p}(0) = \mathbf{p}_L \tag{82}$$

$$0 = \mathbf{z}(t) - \mathbf{I}_{\text{irr}} \cdot \mathbf{K} \cdot \hat{\mathbf{u}}(\mathbf{x}, \mathbf{p}) \tag{83}$$

$$\mathbf{y}(t) = \mathbf{x}_{\text{sel}}(t) \tag{84}$$

$$\mathbf{x}(t) \geq 0, \mathbf{z}(t) \geq 0 \tag{85}$$

with $\mathbf{p}$ containing both the parameters in $\mathbf{P}_{\text{u}}$ and the two $k_{\text{l}}a$ values, $\mathbf{f}_{\text{meta}}$, $\mathbf{f}_{\text{diss}}$ and $\mathbf{f}_{\text{head}}$ as in Equations (64) to (66), $\hat{\mathbf{u}}(\mathbf{x}, \mathbf{p})$ as defined in Equation (77), and $\mathbf{x}_{\text{sel}}$ as defined in Table 6.

The MHE horizon was chosen as 20, i.e., 20 measurement intervals and thus 21 measurement points in the horizon. The $\mathbf{V}$ and $\mathbf{W}_{\text{x}}$ matrices are set up with the noise standard deviations as defined in Table 6. The parameter noise standard deviation was chosen to be 0.0001, based on the considerations presented in the previous section. Initial guesses have to be made for the arrival cost parameters $\bar{\mathbf{x}}_0^{\text{c}}$ and $\bar{\mathbf{P}}_0$. The state estimate is initialized at the measurement at $t = 0$, except for the states which are not measured. These are initialized at 0.01 for the dissolved carbon dioxide concentration, 0.25 for the dissolved oxygen concentration, and 10.0 for the PDO concentration. The initial parameter estimates were set at 0.01 for all parameters in $\mathbf{P}_{\text{u}}$

Journal homepage: http://www.journals.elsevier.com/journal-of-process-control
Original file available at: doi:10.1016/j.jprocont.2015.08.014

and 75.0 for both transfer coefficients. Finally, the standard deviations to be used in $\bar{\mathbf{P}}_0$ are chosen equal to the measurement standard deviation for the measured states, 0.5 for the unmeasured states, and 1.0 for all parameters.

3.2.3. Flux and state estimation results

Because of memory problems due to the large amount of parameter states in the problem with a free $\mathbf{K}$ matrix, this problem could not be solved for this larger case study. To still have an indication whether the choice of a specific $\mathbf{K}$ matrix can have an influence on the estimates, the estimations are performed with two different $\mathbf{K}$ matrices, one with fluxes 45 to 50 (according to the list of reactions shown in the Supplementary data) as free fluxes, and another with fluxes 4, 9, 10, 12, 24 and 39 as free fluxes. This last basis was chosen randomly. The results for the fluxes, for both bases, are plotted in Figure 15, and the results for the states are displayed in Figure 16 for the first basis and in Figure 17 for the second basis. In the figure for the fluxes, only the free fluxes according to the first basis (fluxes 45 to 50) are shown for comparison, all other fluxes can be calculated by left multiplying the free flux vector with the null space basis matrix $\mathbf{K}$.

From these figures, it is clear that also for a larger network, the described methodology obtains accurate estimates for both states and fluxes. For the fluxes, there is always an initial period where the estimates are off, probably due to the inaccurate initial guesses, but after this initial period, most fluxes are practically perfect. For fluxes 45 and 48, which are the PDO production and glucose uptake fluxes, respectively, there is a small delay visible between the true process and the estimates. In the first case, this is probably due to

Journal homepage: http://www.journals.elsevier.com/journal-of-process-control
Original file available at: doi:10.1016/j.jprocont.2015.08.014

the fact that the PDO concentration is not measured. For the glucose flux, no direct explanation can be given. One possible problem could be that the parameter noise levels are the same for all parameters, while in reality, the magnitudes of these parameters can differ significantly, requiring distinct values for each parameter. This problem could be solved by devising a strategy where the parameter noise is estimated from the data in one iteration of the MHE procedure, and used in the next one.

Also for the states, the estimates are very accurate, even for the unmeasured PDO concentration. For the dissolved oxygen and carbon dioxide concentration, however, the correct profile with a small shift downward is observed. Because these are not measured, the general profile can be estimated, but the magnitude of this profile is not accurately identifiable. For the acetate concentration, which is zero over the full experimental horizon in the real process, a very small positive value is estimated everywhere. This happens because the measurement for this concentration is not exactly zero.

Finally, the choice of the $\mathbf{K}$ matrix does not have a big influence on the estimates when using the current methodology, as opposed to the spline-based offline dMFA algorithm in [13]. For both choices of the $\mathbf{K}$ matrix, the results are practically identical over all fluxes, states and time. This indicates that it is not necessary to optimize the $\mathbf{K}$ matrix in the current methodology, which was at present not practically feasible from a computational perspective as well.

3.2.4. Computational results

For both **K** matrices, one iteration of the MHE procedure, including the solution of the dynamic optimization problem and the calculation of the arrival cost parameters for the next iteration, took 1.8 seconds on average on the same CPU as in the small-scale case study. This is, in light of the general bioprocess sampling frequencies, very fast, and thus no big problems on the computational level can be expected when implementing this methodology in a real-life process. If extra speed is necessary, however, parallellization can offer improvements. Although, as opposed to the offline dMFA algorithm, there is no parallel structure in the current methodology, parallellization on a lower level, i.e., on the level of the nonlinear programming solver IPOPT, can dramatically speed up the computations, as shown in [34]. Because of these possible improvements, and the fact that there is still some room on the computational side because of the slow processes and low sampling frequencies in biotechnology, the methodology should scale well to large-scale networks, as long as the number of free fluxes in the network does not increase dramatically. A large increase in the number of free fluxes would mean an even larger increase in the total number of parameters to be estimated, possibly resulting in identifiability problems. In that case, a possible solution could be to integrate a model reduction procedure in the different MHE problems.

4. Conclusions

In this contribution, a methodology for online estimation of metabolic fluxes and concentrations based on dMFA, black-box flux expressions, and

MHE, was presented, with applications on both a small-scale case study and a more realistic, medium-scale network. Two black-box flux models were compared to each other, and to two extremes regarding the integration of mechanistic knowledge into the flux model: the true process model and a model that does not contain any dependency on the states. An important parameter of this methodology is the level of the process noise corresponding to the estimated flux parameters. The effect of this parameter on the estimation performance was assessed. A linear black-box flux model was identified as the best performing model. The performance of the MHE procedure with this linear flux model was then studied on a more realistic case study with a network for *E. coli*, and found to be satisfactory.

For this larger case study, the effect of the choice of the $\mathbf{K}$ matrix on the performance was also determined, and this was found to be far less significant than in the offline flux estimation methodology described in [13]. Based on these results, the described methodology is determined to be a computationally feasible methodology for online estimation of fluxes in metabolic reaction networks. The knowledge of these fluxes can be of large importance in online monitoring and control applications, making it possible to devise advanced control strategies based on the estimated fluxes. Furthermore, the methodology is highly transferable between microorganisms or strains. This enables researchers to quickly define a monitoring and control strategy for their specific bioprocess based on a metabolic reaction network of the organism at hand.

Acknowledgments

Dominique Vercammen has a PhD grant of the Institute for the Promotion of Innovation through Science and Technology in Flanders (IWT-Vlaanderen). Jan Van Impe holds the chair Safety Engineering sponsored by the Belgian chemistry and life sciences federation essenscia. The research was supported by: PFV/10/002 (OPTEC), FWO KAN2013 1.5.189.13, FWO-G.0930.13 and BelSPO: IAP VII/19 (DYSCO).

References

[1] C. Komives, R. S. Parker, Bioreactor state estimation and control, Curr. Opin. Biotech. 14 (2003) 468–474.

[2] D. Dochain, State and parameter estimation in chemical and biochemical processes: a tutorial, J. Process Contr. 13(8) (2003) 801–818.

[3] G. Goffaux, A. Vande Wouwer, Bioprocess state estimation: some classical and less classical approaches, in: T. Meurer, K. Graichen, E. D. Gilles (Eds.), Control and Observer Design for Nonlinear Finite and Infinite Dimensional Systems, Lecture Notes in Control and Information Science, Vol. 322, Springer, 2005, pp. 111–128.

[4] P. Bogaerts, A. Vande Wouwer, Parameter identification for state estimation - application to bioprocess software sensors, Chem. Eng. Sci. 59 (2004) 2465–2476.

[5] A. Vargas, J. A. Moreno, A. Vande Wouwer, A weighted variable gain

super-twisting observer for the estimation of kinetic rates in biological systems, J. Process Contr. 24(6) (2014) 957–965.

[6] H. De Battista, J. Picó, F. Garelli, A. Vignoni, Specific growth rate estimation in (fed-)batch bioreactors using second-order sliding observers, J. Process Contr. 21(7) (2011) 1049–1055.

[7] A. U. Raghunathan, J. R. Pérez-Correa, E. Agosin, L. T. Biegler, Parameter estimation in metabolic flux balance models for batch fermentationformulation & solution using differential variational inequalities (DVIs), Anal. Oper. Res. 148(1) (2006) 251–270.

[8] W. Dai, D. P. Word, J. Hahn, Modeling and dynamic optimization of fuel-grade ethanol fermentation using fed-batch process, Control Eng. Pract. 22 (2014) 231–241.

[9] W. Wiechert, 13C metabolic flux analysis, Metab. Eng. 3(3) (2001) 195–206.

[10] M. R. Antoniewicz, Dynamic metabolic flux analysis - tools for probing transient states of metabolic networks, Curr. Opin. Biotech. 24 (2013) 973–978.

[11] J. Dias, F. Pardelha, M. Eusebio, M. A. M. Reis, R. Oliveira, On-line adaptive metabolic flux analysis: Application to PHB production by mixed microbial cultures, Biotechnol. Progr. 25(2) (2009) 390–398.

[12] M. L. Fazenda, J. M. L. Dias, L. M. Harvey, A. Nordon, R. Edrada-Ebel, D. LittleJohn, B. McNeil, Towards better understanding of an

industrial cell factory: investigating the feasibility of real-time metabolic flux analysis in *Pichia pastoris*, Microb. Cell Fact. 12 (2013) 51.

[13] D. Vercammen, F. Logist, J. Van Impe, Dynamic estimation of specific fluxes in metabolic networks using non-linear dynamic optimization, BMC Syst. Biol. 8 (2014) 132.

[14] G. Goffaux, M. Perrier, M. Cloutier, Cell energy metabolism: A constrained ensemble Kalman filter, in: Proceedings of the 18th IFAC world congress: Milano, Italy, International Federation of Automatic Control, 2011, pp. 8391–8396.

[15] M. Kawohl, T. Heine, R. King, Model based estimation and optimal control of fed-batch fermentation processes for the production of antibiotics, Chem. Eng. Process.: Process Intensif. 46(11) (2007) 1223–1241.

[16] H. L. T. Lee, P. Boccazzi, N. Gorret, R. J. Ram, A. J. Sinskey, In situ bioprocess monitoring of *Escherichia coli* bioreactions using Raman spectroscopy, Vib. Spectrosc. 35 (2004) 131–137.

[17] S. Beutel, S. Henkel, In situ sensor techniques in modern bioprocess monitoring, Appl. Microbiol. Biot. 91(6) (2011) 1493–1505.

[18] L. W. Forman, B. D. Thomas, F. S. Jacobson, On-line monitoring and control of fermentation processes by flow-injection analysis, Anal. Chim. Acta 249(1) (1991) 101–111.

[19] J. F. Van Impe, D. Vercammen, E. Van Derlinden, Toward a next generation of predictive models: A systems biology primer, Food Control 29(2) (2013) 336–42.

[20] J. E. Haag, A. Vande Wouwer, M. Remy, A general model of reaction kinetics in biological systems, Bioproc. Biosyst. Eng. 27(5) (2005) 303–309.

[21] D. G. Robertson, J. H. Lee, J. B. Rawlings, A moving horizon-based approach to least squares estimation, AIChE J. 42(8) (1996) 2209.

[22] P. Kühl, M. Diehl, T. Kraus, J. P. Schlöder, H. G. Bock, A real-time algorithm for moving horizon state and parameter estimation, Comput. Chem. Eng. 35 (2011) 71–83.

[23] M. J. Tenny, J. B. Rawlings, Efficient moving horizon estimation and nonlinear model predictive control, in: Proceedings of the American Control Conference: Anchorage, USA, 2002, pp. 4475–4480.

[24] R. Lopez-Negrete, S. C. Patwardhan, L. T. Biegler, Constrained particle filter approach to approximate the arrival cost in moving horizon estimation, J. Process Contr. 21(6) (2011) 909–919.

[25] S. Ungarala, Computing arrival cost parameters in moving horizon estimation using sampling based filters, J. Process Contr. 19(9) (2009) 1576–1588.

[26] C. C. Qu, J. Hahn, Computation of arrival cost for moving horizon estimation via unscented Kalman filtering, J. Process Contr. 19(2) (2009) 358–363.

[27] D. Vercammen, F. Logist, J. Van Impe, Estimation of specific fluxes in metabolic networks using non-linear dynamic optimization, in:

J. Klemes, F. Friedler, P. Mizsey (Eds.), Proceedings of the 24th European Symposium of Computer Aided Process Engineering, Budapest, Hungary, 2014, pp. 289–294.

[28] A. Wächter, L. T. Biegler, On the implementation of a primal-dual interior point filter line search algorithm for large-scale nonlinear programming, Math. Program. 106(1) (2006) 25–57.

[29] J. Andersson, J. Åkesson, M. Diehl, CasADi – A symbolic package for automatic differentiation and optimal control, in: S. Forth, P. Hovland, E. Phipps, J. Utke, A. Walther (Eds.), Recent Advances in Algorithmic Differentiation, Vol. 87 of Lecture Notes in Computational Science and Engineering, Springer, Berlin Heidelberg, 2012, pp. 297–307.

[30] J. D. Orth, R. M. T. Fleming, B. O. Palsson, Reconstruction and use of microbial metabolic networks: the core *Escherichia coli* metabolic model as an educational guide, in: A. Böck, R. Curtis III, J. Kaper, P. Karp, F. Neidhardt, T. Nyström, J. Slauch, C. Squires, D. Ussery (Eds.), EcoSal-Escherichia coli and Salmonella: Cellular and molecular biology, ASM Press, Washington D.C., USA, 2009, pp. 56–99.

[31] R. W. Leighty, M. R. Antoniewicz, Dynamic metabolic flux analysis (DMFA): a framework for determining fluxes at metabolic non-steady state, Metab. Eng. 13(6) (2011) 745–755.

[32] R. Mahadevan, J. Edwards, F. Doyle, Dynamic flux balance analysis of diauxic growth in *Escherichia coli*, Biophys. J. 83 (3) (2002) 1331–1340.

[33] A. L. Meadows, R. Karnik, H. Lam, S. Forestell, B. Snedecor, Application of dynamic flux balance analysis to an industrial *Escherichia coli* fermentation, Metab. Eng. 12 (2) (2010) 150–160.

[34] V. M. Zavala, C. D. Laird, L. T. Biegler, Interior-point decomposition approaches for parallel solution of large-scale nonlinear parameter estimation problems, Chem. Eng. Sci. 63 (2008) 4834–4845.

Figures

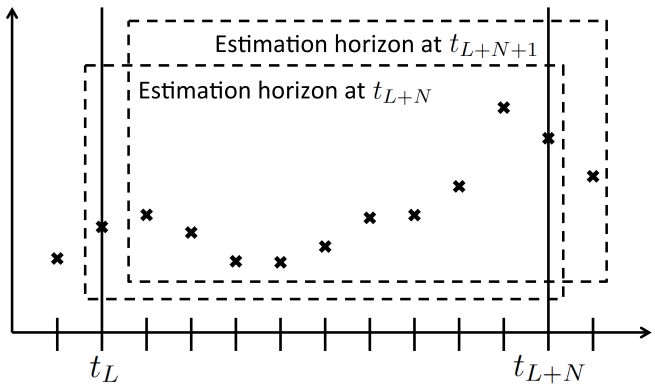
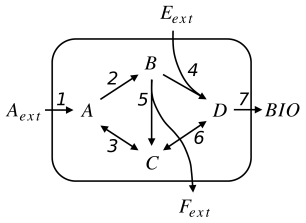


Figure 1: **Moving horizon estimation.** Graphical representation of the estimation horizons of two consecutive MHE problems, at t_{L+N} and t_{L+N+1}. The horizon length N is in this case equal to 10.

Journal homepage: http://www.journals.elsevier.com/journal-of-process-control
Original file available at: doi:10.1016/j.jprocont.2015.08.014

Postprint version of paper published in Journal of Process Control 2016, vol. 37, pages 1-20.
The content is identical to the published paper, but without the final typesetting by the publisher.



$$\mathbf{S}_{\text{int}} = \begin{bmatrix} 1 & -1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 & -2 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 2 & 0 & 1 & -1 \end{bmatrix}$$

$$\mathbf{S}_{\text{e}} = \begin{bmatrix} -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

$$\mathbf{I}_{\text{irr}} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

$$\mathbf{K}_{145} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 2 \\ 1 & -1 & -2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & -1 & -1 \\ 1 & 1 & -1 \end{bmatrix}$$

Figure 2: **Small-scale case study network and corresponding matrices.** Metabolic reaction network for the small-scale case study (top left), along with the intracellular and combined extracellular and biomass stoichiometric matrices and irreversibility matrix (right), and the null space basis matrix corresponding to free fluxes 1, 4 and 5 (bottom left).

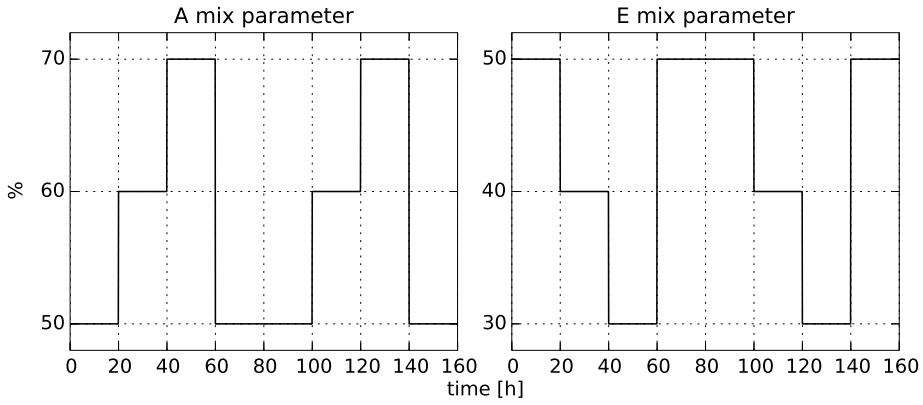


Figure 3: **Small-scale case study input profiles.** The time profiles for the controls r_A and r_E in the small-scale case study.

Journal homepage: http://www.journals.elsevier.com/journal-of-process-control
Original file available at: doi:10.1016/j.jprocont.2015.08.014

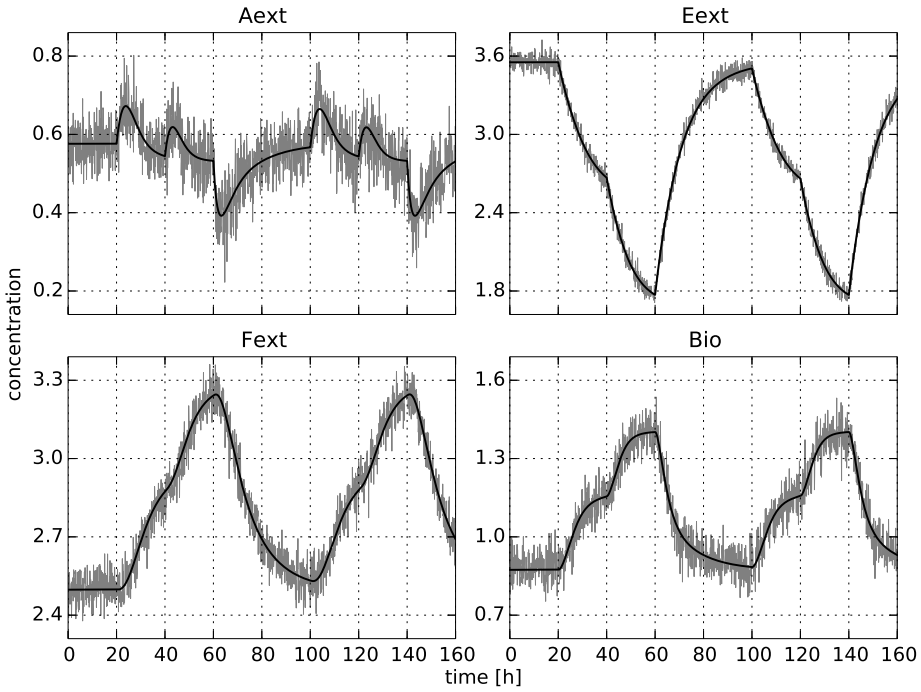


Figure 4: **Small-scale case study measurements.** The simulated measurements for the different outputs in the small-scale case study.

Postprint version of paper published in Journal of Process Control 2016, vol. 37, pages 1-20.
The content is identical to the published paper, but without the final typesetting by the publisher.

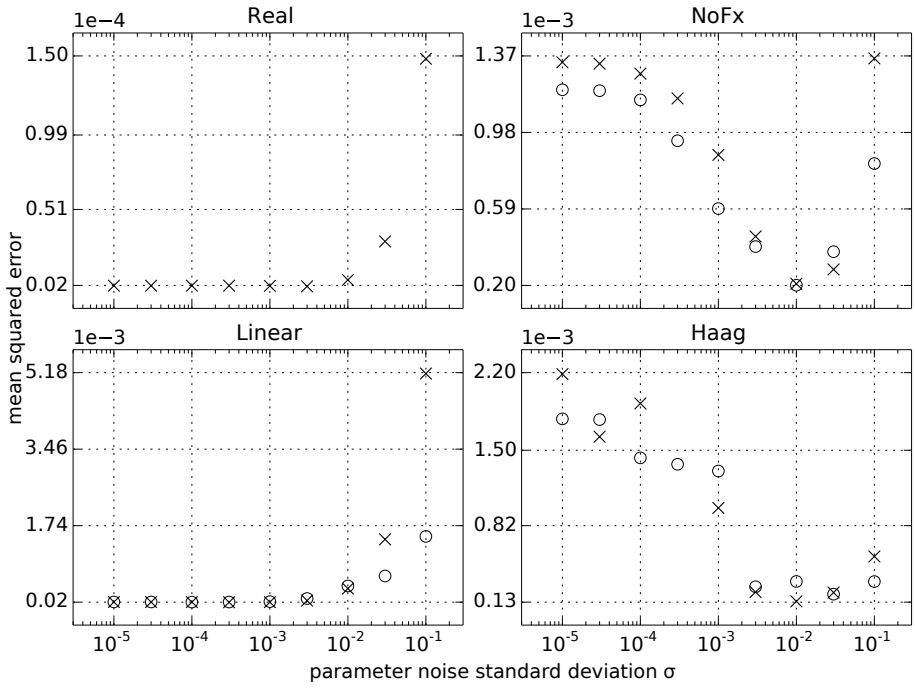


Figure 5: **Small-scale case study flux errors.** Mean squared errors between simulated and estimated fluxes for the different scenarios in the small-scale case study. Crosses indicate results for the estimations with fixed $\mathbf{K}$ matrix, circles for the estimations with a free $\mathbf{K}$ matrix.

Journal homepage: http://www.journals.elsevier.com/journal-of-process-control
Original file available at: doi:10.1016/j.jprocont.2015.08.014

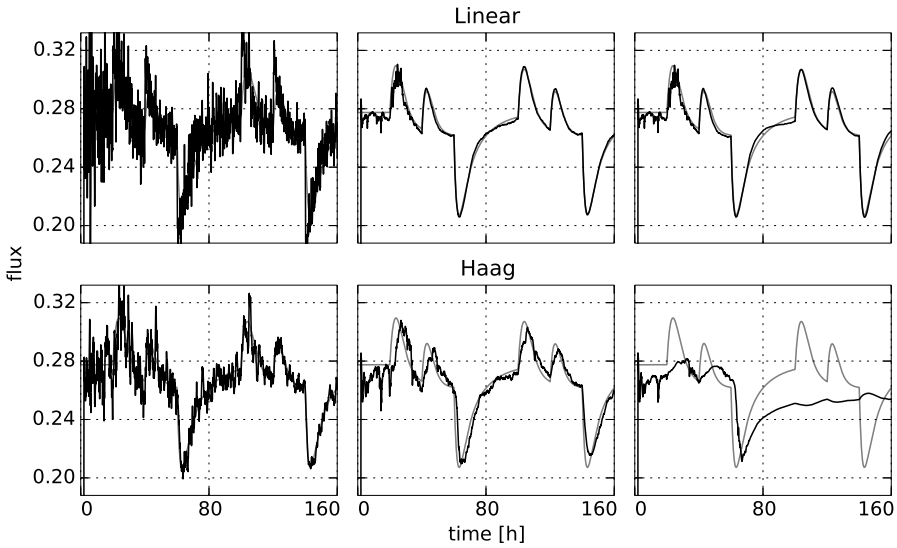


Figure 6: **Small-scale case study flux results.** Simulated (in gray) and estimated fluxes (in black) for the *Linear* and *Haag* models with fixed **K** matrix. The shown fluxes are for reaction 1 in the network. The figures on the left correspond to a parameter noise standard deviation of 0.1, the ones on the right to a parameter noise level of 10^{-5}, and the ones in the middle to the optimum value as defined in Table 3.

Postprint version of paper published in Journal of Process Control 2016, vol. 37, pages 1-20.
The content is identical to the published paper, but without the final typesetting by the publisher.

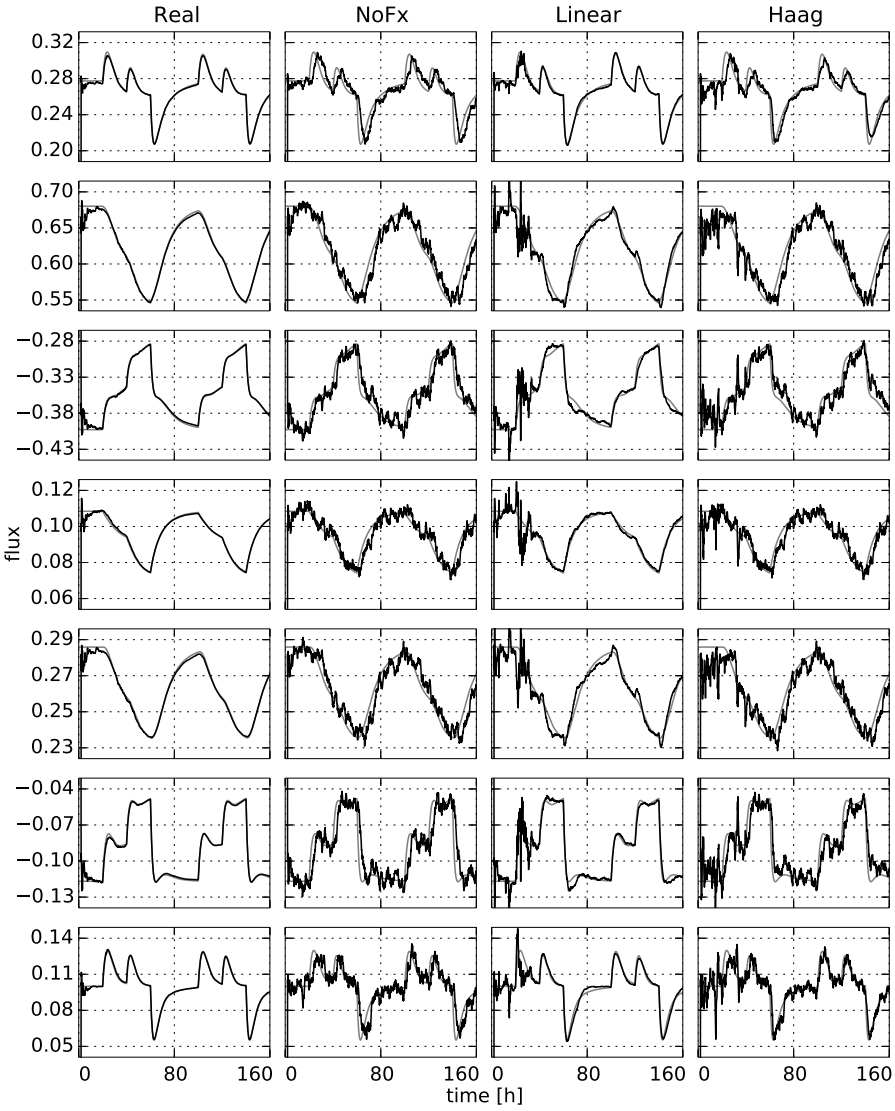


Figure 7: **Small-scale case study flux results.** Simulated (in gray) and estimated (in black) fluxes in the different scenarios with a fixed $\mathbf{K}$ matrix, for the optimal parameter noise level corresponding to the case, going from the flux for reaction 1 at the top to the flux for reaction 7 at the bottom.

Journal homepage: http://www.journals.elsevier.com/journal-of-process-control
Original file available at: doi:10.1016/j.jprocont.2015.08.014

Postprint version of paper published in Journal of Process Control 2016, vol. 37, pages 1-20.
The content is identical to the published paper, but without the final typesetting by the publisher.

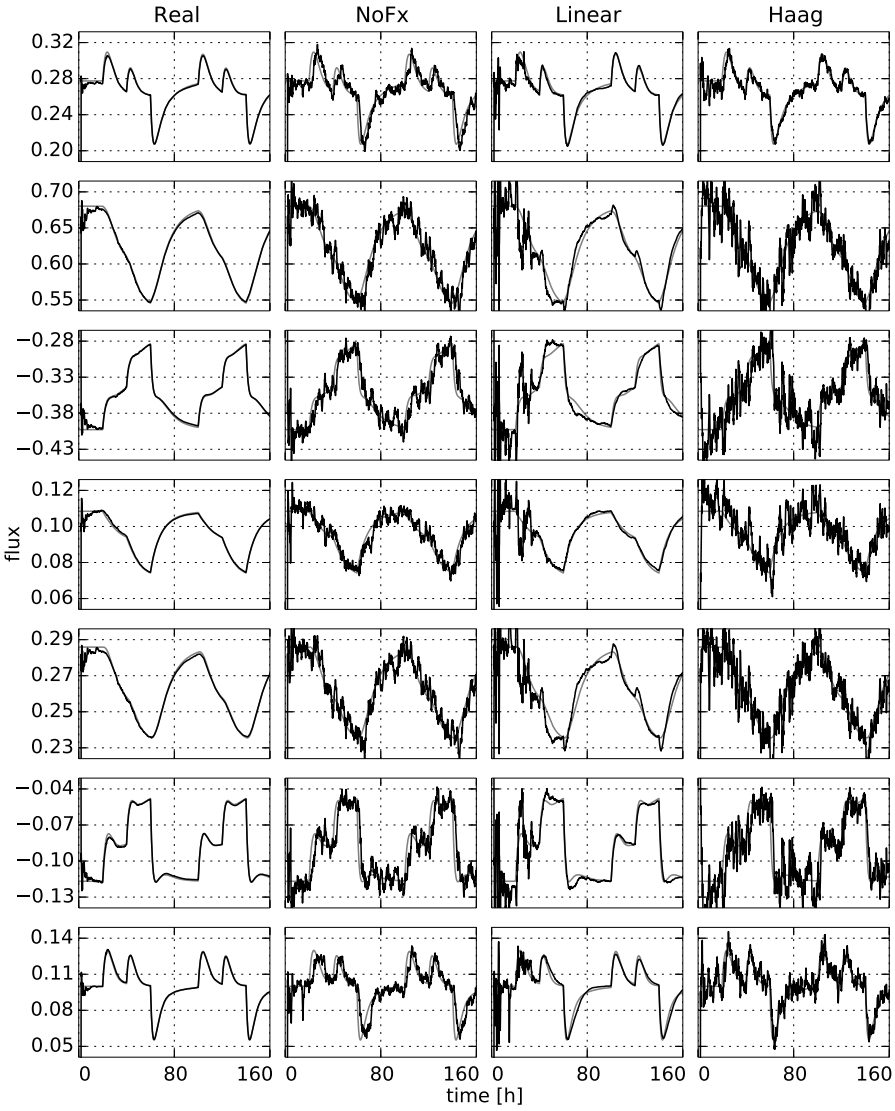


Figure 8: **Small-scale case study flux results.** Simulated (in gray) and estimated (in black) fluxes in the different scenarios with a free $\mathbf{K}$ matrix, for the optimal parameter noise level corresponding to the case, going from the flux for reaction 1 at the top to the flux for reaction 7 at the bottom.

Journal homepage: http://www.journals.elsevier.com/journal-of-process-control
Original file available at: doi:10.1016/j.jprocont.2015.08.014

Postprint version of paper published in Journal of Process Control 2016, vol. 37, pages 1-20.
The content is identical to the published paper, but without the final typesetting by the publisher.

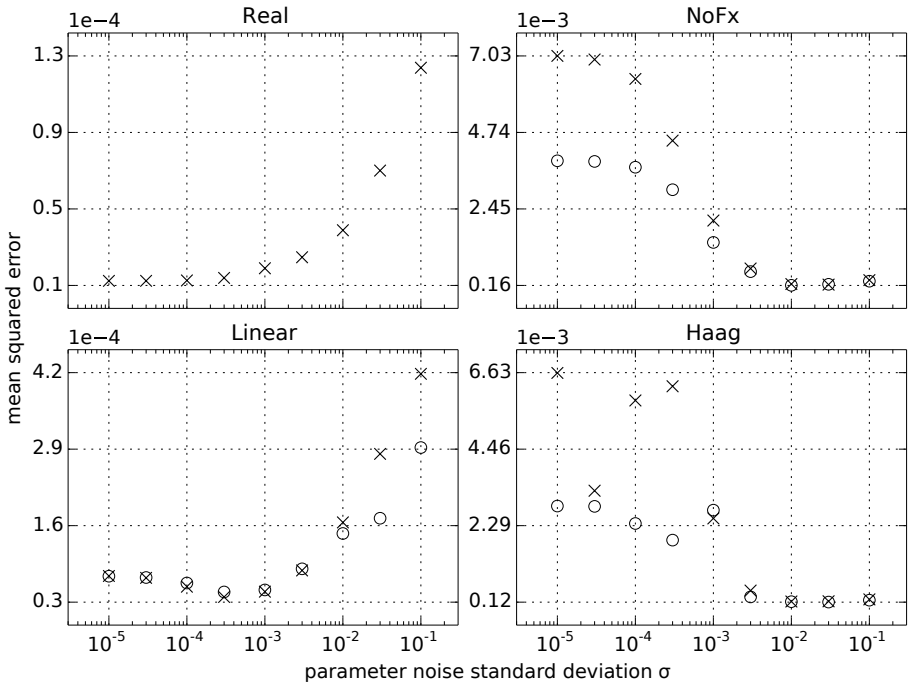


Figure 9: **Small-scale case study state errors.** Mean squared errors between simulated and estimated states for the different scenarios in the small-scale case study. Crosses indicate results for the estimations with fixed $\mathbf{K}$ matrix, circles for the estimations with a free $\mathbf{K}$ matrix.

Journal homepage: http://www.journals.elsevier.com/journal-of-process-control
Original file available at: doi:10.1016/j.jprocont.2015.08.014

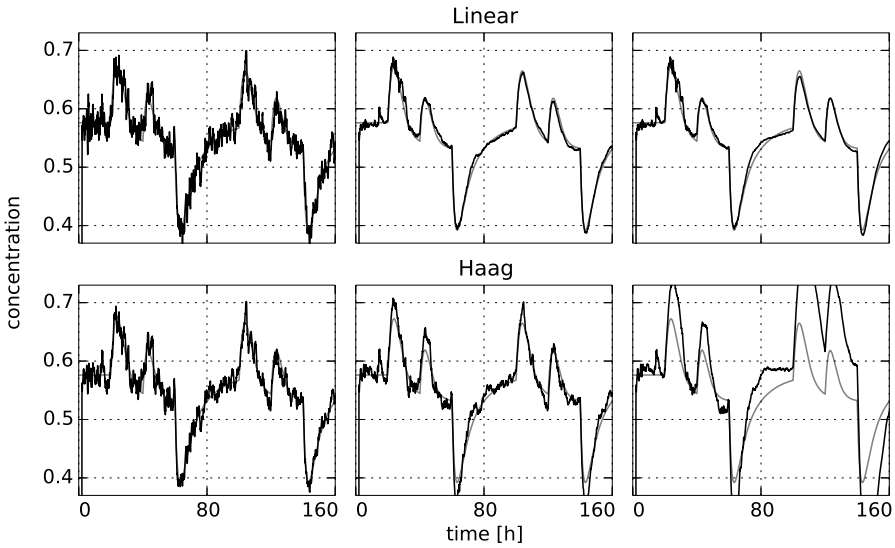


Figure 10: **Small-scale case study state results.** Simulated (in gray) and estimated states (in black) for the *Linear* and *Haag* models with fixed **K** matrix. The time profiles for the concentration of Aext are shown. The figures on the left correspond to a parameter noise standard deviation of 0.1, the ones on the right to a parameter noise level of 10^{-5}, and the ones in the middle to the optimum value as defined in Table 4.

Journal homepage: http://www.journals.elsevier.com/journal-of-process-control
Original file available at: doi:10.1016/j.jprocont.2015.08.014

Postprint version of paper published in Journal of Process Control 2016, vol. 37, pages 1-20.
The content is identical to the published paper, but without the final typesetting by the publisher.

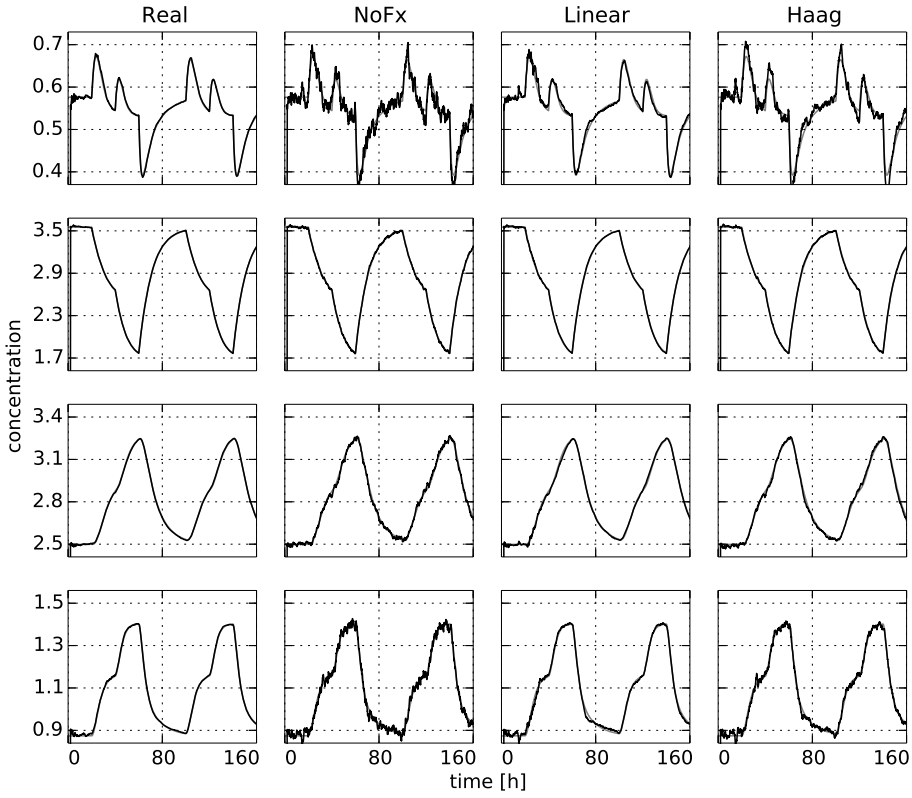


Figure 11: **Small-scale case study state results.** Simulated (in gray) and estimated (in black) states in the different scenarios with a fixed **K** matrix, for the optimal parameter noise level corresponding to the case, for the concentrations of Aext, Eext, Fext and Biomass from top to bottom.

Journal homepage: http://www.journals.elsevier.com/journal-of-process-control
Original file available at: doi:10.1016/j.jprocont.2015.08.014

Postprint version of paper published in Journal of Process Control 2016, vol. 37, pages 1-20.
The content is identical to the published paper, but without the final typesetting by the publisher.

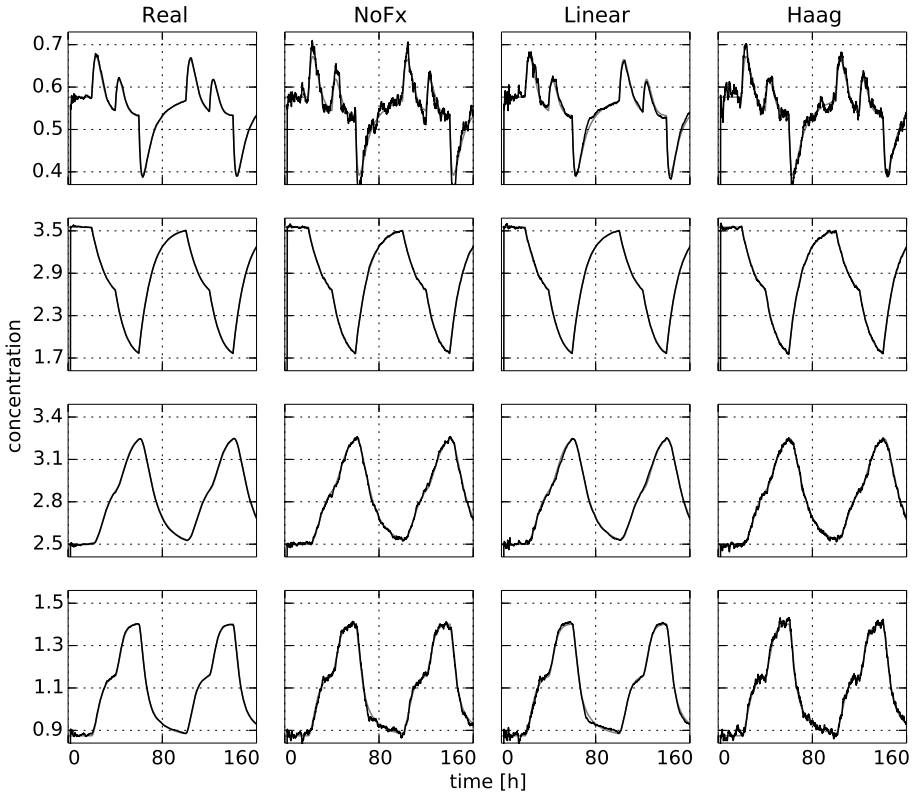


Figure 12: **Small-scale case study state results.** Simulated (in gray) and estimated (in black) states in the different scenarios with a free **K** matrix, for the optimal parameter noise level corresponding to the case, for the concentrations of Aext, Eext, Fext and Biomass from top to bottom.

Journal homepage: http://www.journals.elsevier.com/journal-of-process-control
Original file available at: doi:10.1016/j.jprocont.2015.08.014

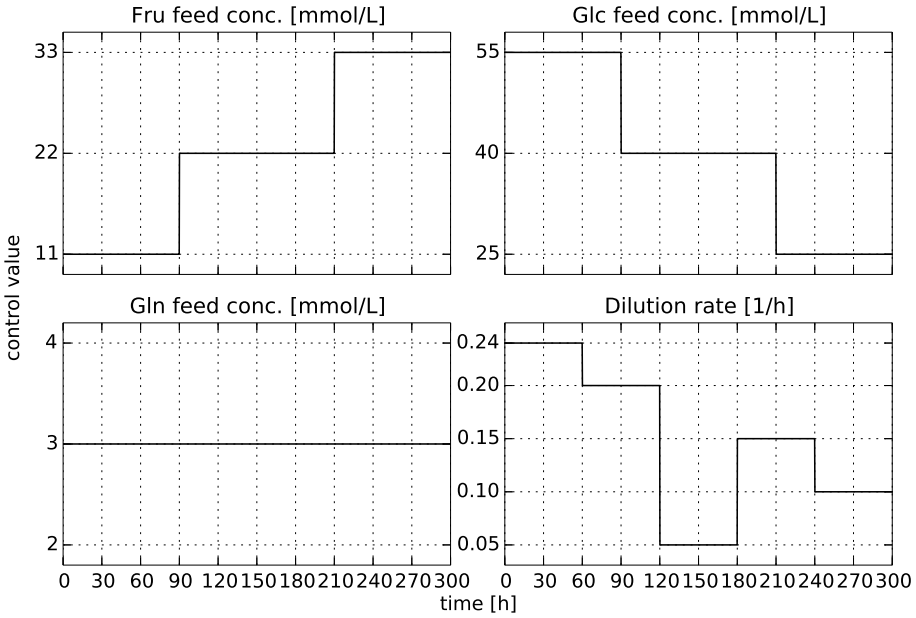


Figure 13: **Medium-scale case study input profiles.** The time profiles for the controls in the medium-scale case study for *E. coli*.

Postprint version of paper published in Journal of Process Control 2016, vol. 37, pages 1-20.
The content is identical to the published paper, but without the final typesetting by the publisher.

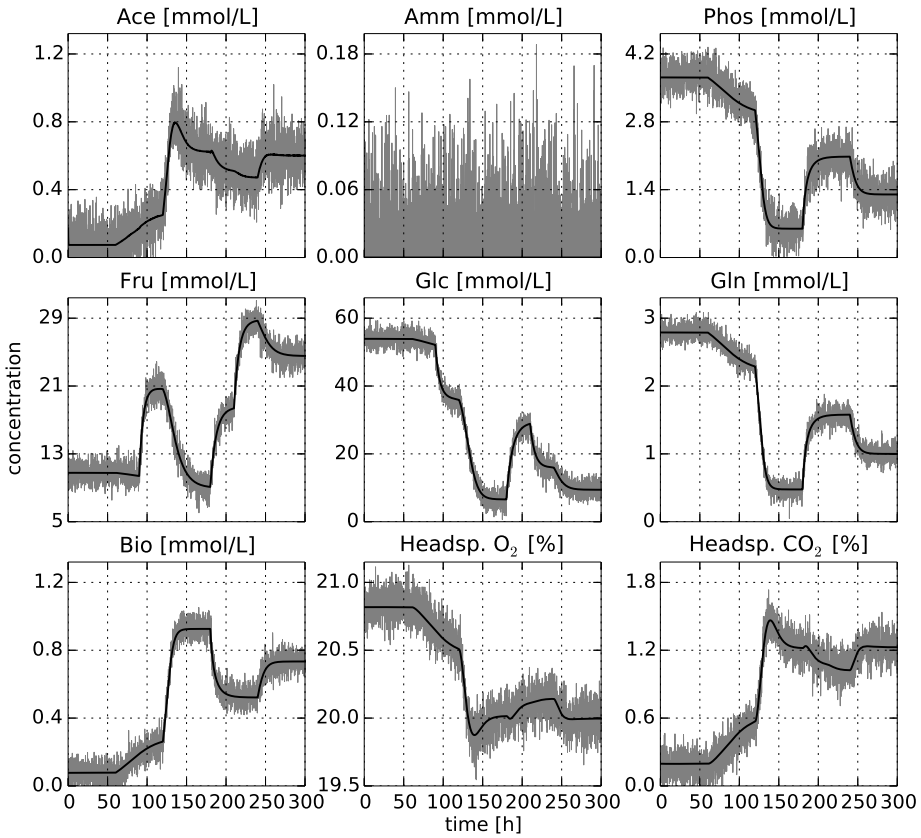


Figure 14: **Medium-scale case study measurements.** The simulated measurements for the different outputs in the medium-scale case study for *E. coli*.

Journal homepage: http://www.journals.elsevier.com/journal-of-process-control
Original file available at: doi:10.1016/j.jprocont.2015.08.014

Postprint version of paper published in Journal of Process Control 2016, vol. 37, pages 1-20.
The content is identical to the published paper, but without the final typesetting by the publisher.

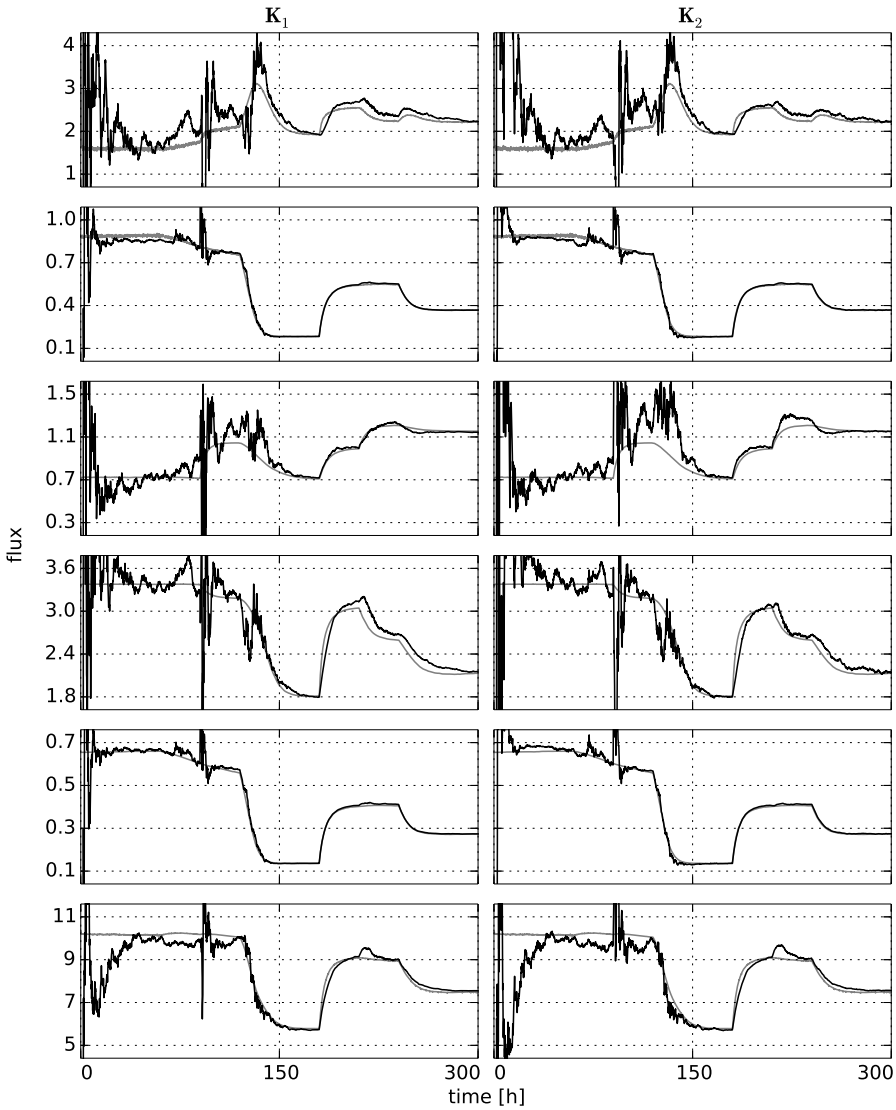


Figure 15: **Medium-scale case study flux results.** Simulated (in gray) and estimated (in black) fluxes for the medium-scale case study for *E. coli*, for basis 1 on the left and for basis 2 on the right, for fluxes 45 to 50 from top to bottom.

Journal homepage: http://www.journals.elsevier.com/journal-of-process-control
Original file available at: doi:10.1016/j.jprocont.2015.08.014

Postprint version of paper published in Journal of Process Control 2016, vol. 37, pages 1-20.
The content is identical to the published paper, but without the final typesetting by the publisher.

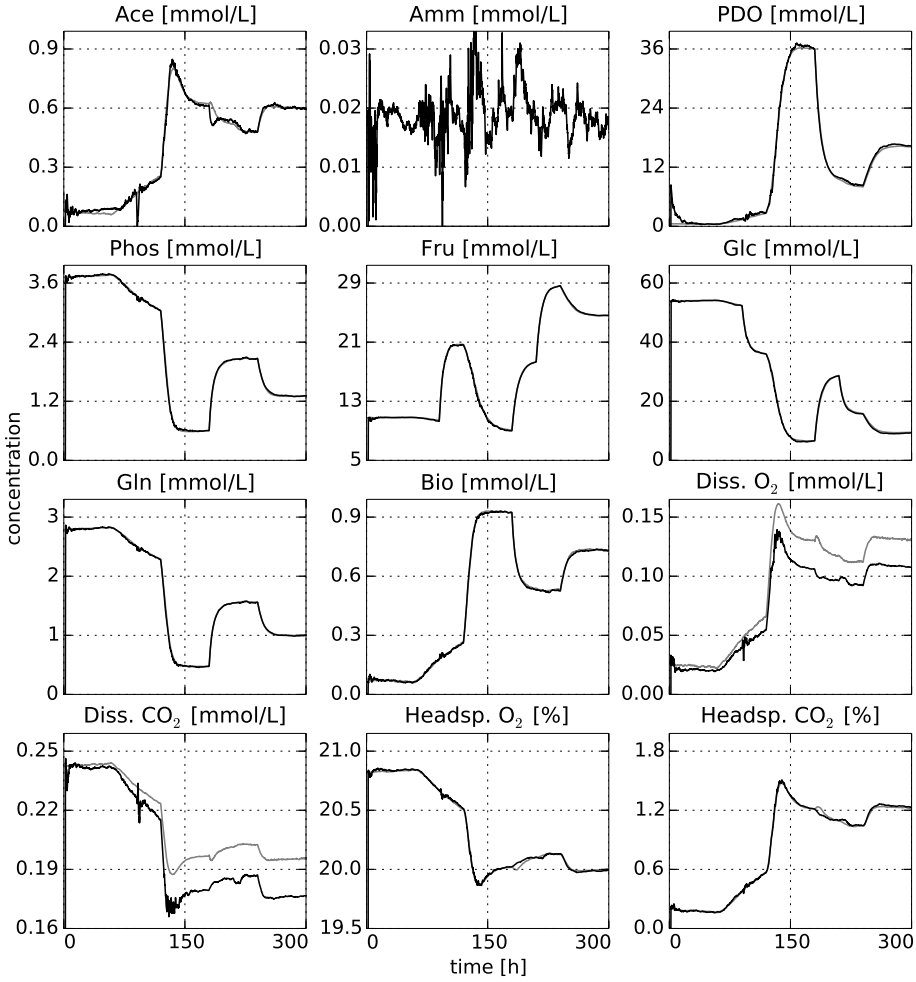


Figure 16: **Medium-scale case study state results.** Simulated (in gray) and estimated (in black) states for the medium-scale case study for *E. coli* with basis 1.

Journal homepage: http://www.journals.elsevier.com/journal-of-process-control
Original file available at: doi:10.1016/j.jprocont.2015.08.014

Postprint version of paper published in Journal of Process Control 2016, vol. 37, pages 1-20.
The content is identical to the published paper, but without the final typesetting by the publisher.

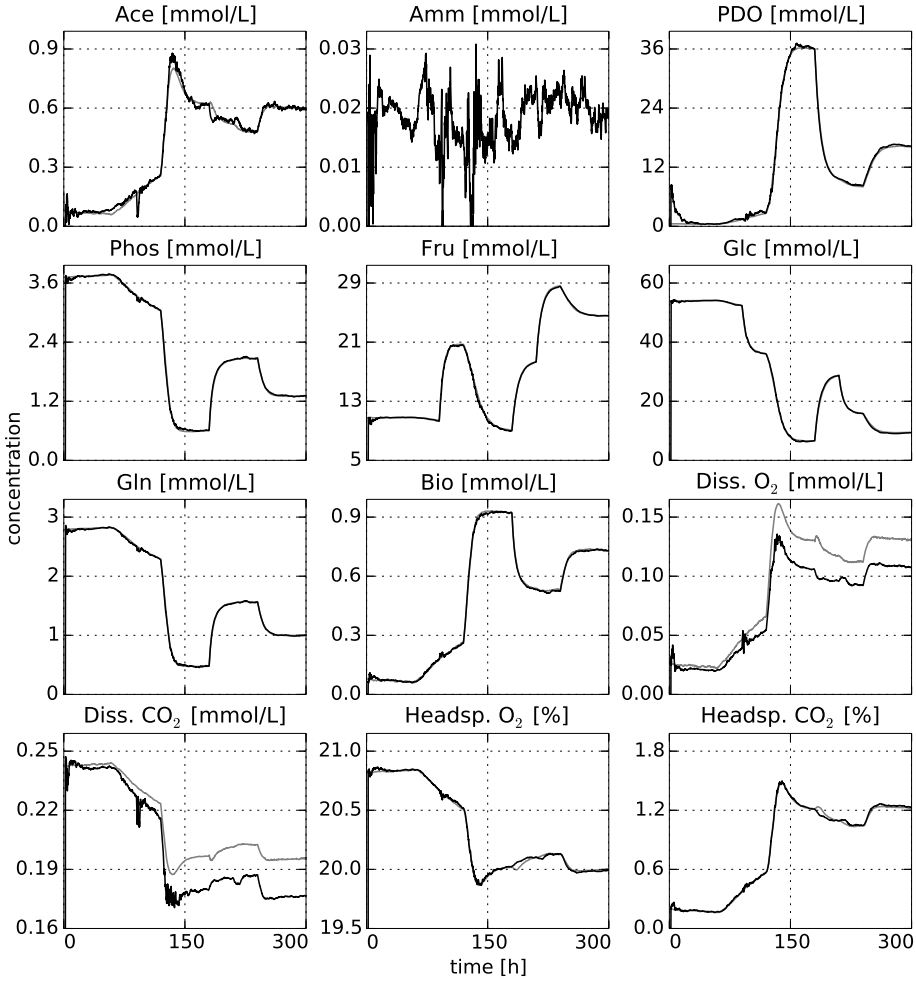


Figure 17: **Medium-scale case study state results.** Simulated (in gray) and estimated (in black) states for the medium-scale case study for *E. coli* with basis 2.

Journal homepage: http://www.journals.elsevier.com/journal-of-process-control
Original file available at: doi:10.1016/j.jprocont.2015.08.014

Tables

Table 1: **Small-scale case study state overview.** Overview of the states in the small-scale case study, with initial concentrations in the simulation, concentrations in the feed, process noises and measurement noises. The $*$ indicate that these feed concentrations in the final mix are controlled via the mix parameters r_A and r_E. The reported noise values are the standard deviations σ.

Metabolite	Init. conc.	Feed conc.		Proc. noise	Measured	Meas. noise
Aext	0.5760	6.0	*	0.005	✓	0.05
Eext	3.5527	9.0	*	0.005	✓	0.05
Fext	2.4976	0.0		0.005	✓	0.05
Bio	0.8736	0.0		0.005	✓	0.05

Journal homepage: http://www.journals.elsevier.com/journal-of-process-control
Original file available at: doi:10.1016/j.jprocont.2015.08.014

Postprint version of paper published in Journal of Process Control 2016, vol. 37, pages 1-20.
The content is identical to the published paper, but without the final typesetting by the publisher.

Table 2: **Small-scale case study scenarios.** Overview of the different scenarios in the small-scale case study, with parameters to be estimated in each scenario, and initial estimates for these parameters as initial guesses for the arrival cost.

Scenario	Parameters	Number of indep. parameters	Initial estimate $\bar{\mathbf{p}}_0$
Real	$\mathbf{u}_{\max}$	6	1.0 for all $\mathbf{u}_{\max,i}$
	$\mathbf{K}_{\mathrm{M}}$		1.0 for all $\mathbf{K}_{\mathrm{M},i}$
NoFx, $\mathbf{K} = \mathbf{K}_{145}$	$\mathbf{p}_{\mathrm{u}}$	3	0.1 for all $\mathbf{p}_{\mathrm{u},i}$
NoFx, $\mathbf{K}$ free	$\mathbf{p}_{\mathrm{u}}$	3+3	0.1 for all $\mathbf{p}_{\mathrm{u},i}$
	$\mathbf{K}$		$\mathbf{K}_{145}$ for $\mathbf{K}$
Linear, $\mathbf{K} = \mathbf{K}_{145}$	$\mathbf{P}_{\mathrm{u}}$	9	0.01 for all $\mathbf{p}_{\mathrm{u},ij}$
Linear, $\mathbf{K}$ free	$\mathbf{P}_{\mathrm{u}}$	9+3	0.01 for all $\mathbf{P}_{\mathrm{u},ij}$
	$\mathbf{K}$		$\mathbf{K}_{145}$ for $\mathbf{K}$
Haag, $\mathbf{K} = \mathbf{K}_{145}$	$\mathbf{u}_{\max}$	12	0.1 for all $\mathbf{u}_{\max,i}$
	$\mathbf{K}_{\mathrm{H}}$		0.01 for all $\mathbf{K}_{\mathrm{H},ij}$
Haag, $\mathbf{K}$ free	$\mathbf{u}_{\max}$	12+3	0.1 for all $\mathbf{u}_{\max,i}$
	$\mathbf{K}_{\mathrm{H}}$		0.01 for all $\mathbf{K}_{\mathrm{H},ij}$
	$\mathbf{K}$		$\mathbf{K}_{145}$ for $\mathbf{K}$

Journal homepage: http://www.journals.elsevier.com/journal-of-process-control
Original file available at: doi:10.1016/j.jprocont.2015.08.014

Postprint version of paper published in Journal of Process Control 2016, vol. 37, pages 1-20.
The content is identical to the published paper, but without the final typesetting by the publisher.

Table 3: **Small-scale case study parameter noise results.** Numeric results for the minimum points in Figure 5, along with the corresponding optimal parameter noise value.

Scenario	$\mathbf{K} = \mathbf{K}_{145}$		$\mathbf{K}$ free	
	Opt. MSE ($\times 10^{-5}$)	Opt. par. noise	Opt. MSE ($\times 10^{-5}$)	Opt. par. noise
Real	0.18	0.003		
NoFx	20.8	0.01	20.1	0.01
Linear	1.84	0.0003	2.16	0.0001
Haag	13.7	0.01	20.2	0.03

Table 4: **Small-scale case study parameter noise results.** Numeric results for the minimum points in Figure 9, along with the corresponding optimal parameter noise value.

Scenario	$\mathbf{K} = \mathbf{K}_{145}$		$\mathbf{K}$ free	
	Opt. MSE ($\times 10^{-5}$)	Opt. par. noise	Opt. MSE ($\times 10^{-5}$)	Opt. par. noise
Real	1.24	0.000 01		
NoFx	18.4	0.03	16.3	0.01
Linear	3.88	0.0003	4.73	0.0003
Haag	13.8	0.01	12.2	0.03

Journal homepage: http://www.journals.elsevier.com/journal-of-process-control
Original file available at: doi:10.1016/j.jprocont.2015.08.014

Table 5: **Small-scale case study computation times.** Average computation times per MHE iteration for the different scenarios, in milliseconds.

Scenario	Average computation time per MHE iteration [ms]	
	$\mathbf{K} = \mathbf{K}_{145}$	$\mathbf{K}$ free
Real	26.5	
NoFx	17.4	75.8
Linear	40.4	154.7
Haag	68.3	306.3

Journal homepage: http://www.journals.elsevier.com/journal-of-process-control
Original file available at: doi:10.1016/j.jprocont.2015.08.014

Table 6: **Medium-scale case study state overview.** Overview of the states in the medium-scale case study, with initial concentrations in the simulation, concentrations in the feed, process noises and measurement noises. The $*$ indicate that these feed concentrations are controlled. The reported noise values are the standard deviations σ.

Metabolite	Init. conc.	Feed conc.	Proc. noise	Measured	Meas. noise
Metabolite states $\mathbf{x}_{\text{meta}}$					
Ace	0.075	0.0	0.005	✓	0.1
Amm	0.0	0.0	0.0005	✓	0.05
PDO	0.481	0.0	0.05		
Phos	3.716	4.0	0.005	✓	0.25
Fru	10.762	*	0.05	✓	1.0
Glc	53.907	*	0.05	✓	2.0
Gln	2.790	*	0.005	✓	0.1
Bio	0.077	0.0	0.005	✓	0.05
Dissolved states $\mathbf{x}_{\text{diss}}$					
O_2	0.242	0.25	0.0005		
CO_2	0.026	0.01	0.005		
Headspace states $\mathbf{x}_{\text{head}}$					
O_2	0.2082	0.2095	0.005	✓	0.1
CO_2	0.0019	0.0004	0.005	✓	0.1

Table 7: **Medium-scale case study model parameter values.** Model parameter values for the simulation model of the medium-scale case study.

Parameter	Symbol	Value	Units
Diss. oxygen saturation concentration	$\mathbf{x}^*_{\mathrm{diss},\mathrm{O}_2}$	0.25	mmol/L
Diss. carbon dioxide saturation concentration	$\mathbf{x}^*_{\mathrm{diss},\mathrm{CO}_2}$	0.01	mmol/L
Oxygen transfer coefficient	$(k_\mathrm{l}a)_{\mathrm{O}_2}$	100	1/h
Carbon dioxide transfer coefficient	$(k_\mathrm{l}a)_{\mathrm{CO}_2}$	60	1/h
Bioreactor medium volume	V_liq	2	L
Bioreactor headspace volume	V_head	1	L
Reciprocal of ideal gas molar volume	M	40.82	mmol/L
Inlet air flow rate	F	30.0	L/h

Journal homepage: http://www.journals.elsevier.com/journal-of-process-control
Original file available at: doi:10.1016/j.jprocont.2015.08.014